Project 033 Alternative Fuels Test Database Library (Year V)

University of Illinois Urbana-Champaign

Project Lead Investigator

Tonghun Lee
Professor
Mechanical Science & Engineering
University of Illinois at Urbana-Champaign
1206 W. Green St.
Urbana, IL 61801
517-290-8005
tonghun@illinois.edu

University Participants

University of Illinois at Urbana-Champaign

- PI(s): Tonghun Lee, Professor
- FAA Award Number: 13-C-AJFE-UI-026
- Period of Performance: August 15, 2018 to August 14, 2019
- Task(s):
    1. Launch of generation II online database
    2. Preliminary efforts to link the database with JETSCREEN

Project Funding Level

Funding Level: $130,000
Cost Share: Software license support from Reaction Design (ANSYS)

Investigation Team

- Tonghun Lee (Professor, University of Illinois at Urbana-Champaign): Overall research supervision
- Anna Oldani (Graduate Student, University of Illinois at Urbana-Champaign): Compilation of fuel test data and database development

Project Overview

This study seeks to develop a comprehensive and foundational database of current and emerging alternative jet fuels by integrating relevant pre-existing jet fuel data into a common archive that can support scientific research, enhance operational safety, and provide guidelines for the design and certification of new jet fuels. In previous years of this project, efforts were focused on the integration and analysis of pre-existing jet fuel data from various government agencies and individual research groups. Over the last year, we have started to convert all of the compiled data to a new nonstructured query language (NoSQL) format using a JavaScript object notation (JSON) schema, thus allowing the data to be analyzed in a flexible manner using various programming languages. To this end, we have launched the second generation of our online database, which utilizes a new nonrelational database structure. This version will provide interactive analysis functions for users and flexible methods for plotting and downloading data. We have also made significant progress in integrating our database with a similar database in the European JETSCREEN program. Through this process, we have identified new goals, including the need to collect and archive data from fuels used on flights. With these future developments, we hope that the database will one day not only serve as a comprehensive and centralized knowledge base utilized by the aviation community but will also be a resource that can enhance operation efficiency and safety. With the prolific diversification of new jet fuels expected in the near future, particularly of alternative jet fuels, the ability to track critical fuel properties and test data from both research and operation perspectives is expected to be highly valuable for the future of commercial aviation.

Task 1 – Launch of the Generation II Online Database

University of Illinois Urbana-Champaign

Objective(s)

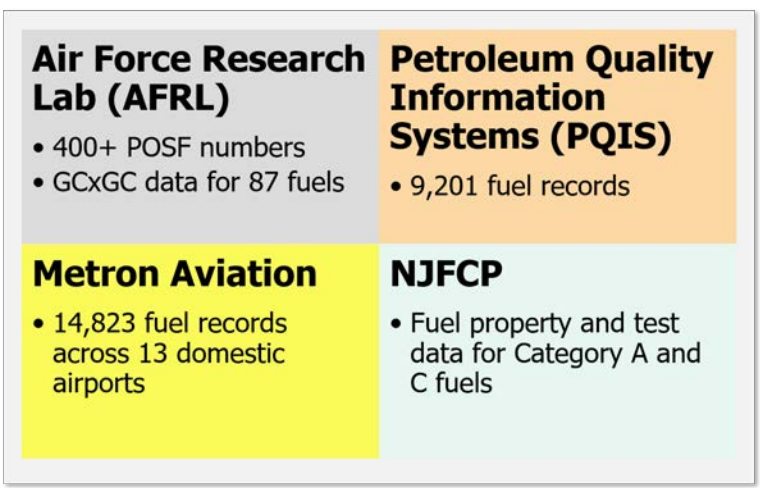
The main objective of this task is to design and launch the generation II Alternative Jet Fuels Test Database online. The goals that guided the first-generation database, namely, establishing a foundational repository of current and emerging information on alternative jet fuels and fuel blends, would be continued in this version. However, the generation II database would exceed our original goal and would be designed using a completely new architecture that allows for flexible analysis and scaling in the future. The new database would utilize a NoSQL format that would be sufficiently flexible to accommodate various data types and that could be easily accessed by any common programming language. This database would also allow large-scale analysis of alternative jet fuel data using advanced algorithms, such as machine learning, and would potentially be linked with other similar databases in the future. The specific goals of this task are as follows:

- Design a nonrelational database and schema for fuel properties and test data
- Convert the entire dataset to the nonrelational JSON (Schema) format
- Build a new generation II database and launch this database online
- Provide interactive analysis capabilities for users

Research Approach

Strategy for the converting the generation II database to NoSQL

The significant task of selecting a database structure and conversion process was completed last year. Currently, the database has grown significantly to house over 25,000 separate fuel records, as shown in the top panel of Figure 1.



Non-Relational (NoSQL)	*Relational (SQL)*
Flexible schema variable data can be easily inserted/altered	**Structured schema** new data must be consistent with predefined structure

Figure 1. Overview of data records in the current database and comparison of relational and nonrelational databases.

The catalogue of data currently available in the database was primarily assembled from four separate sources. The fuels with POSF (AFRL fuel database code) number designations were added from the internal database at the Air Force Research Laboratory (Wright Patterson Air Force Base), which we carefully parsed to identify important data. The second dataset was obtained from the PQIS reports of the Naval Air Systems Command (NAVAIR) and corresponds to a compilation of fuel data geared primarily towards military use. The third set was provided by Metron Aviation, who compiled fuel properties from samples collected at airports through a previous ASCENT project. The dataset resulting from this study proved valuable by providing a landscape of fuels currently used in commercial aviation and guided our future efforts focused on capturing this type of data in real time, as will be discussed below for Task 2. The final dataset was obtained from the National Jet Fuel Combustion Program within ASCENT; for this set, we moved all of the testing data from the KSN to the current database.

Conversion to a NoSQL schema

The database initially housed a large collection of information in many different formats. In year IV, we decided to convert the entire database to a NoSQL structure, primarily because the fuel documents and reports had different formats and varying amounts of data. For example, a dataset for one fuel may include viscosity, density, and heat of combustion measurements, while a dataset for another fuel may only include density. Furthermore, multiple tests may be performed for a single fuel property. A flexible schema would provide a means for storing and retrieving data in a flexible manner and would allow analysis via various programming languages. This change would also enable the incorporation of new information types in the database without major changes to the existing dataset. As an additional benefit of this conversion, the database could be integrated with other similar databases. An effort to perform this type of integration with the European JETSCREEN database will be described in the next section.

After the NoSQL format had been chosen, the selection of a NoSQL service was required. We chose DynamoDB, which is operated by Amazon Web Services (AWS) and has an ongoing contract with the University of Illinois. However, the data would be stored in a nonrelational schema termed JSON, which is language-independent and can be easily interfaced with any other platform should we choose to move away from AWS in the future. The next step involved converting our entire dataset to the JSON format. The JSON format stores data using value pairs, of which the first is a string denoted as a key. However, values can also act as keys pointing to other values, resulting in a complex schema with nested keys. To access data, these specific keys are retrieved through the programming language. This complexity allows for a flexible representation of multiple data types. In this process, we developed an individual JSON format for each test and then wrote automated scripts using Python to convert the data into the nonrelational format. An example of a JSON schema for one of the most complex datasets (GCxGC) is shown in Figure 2.

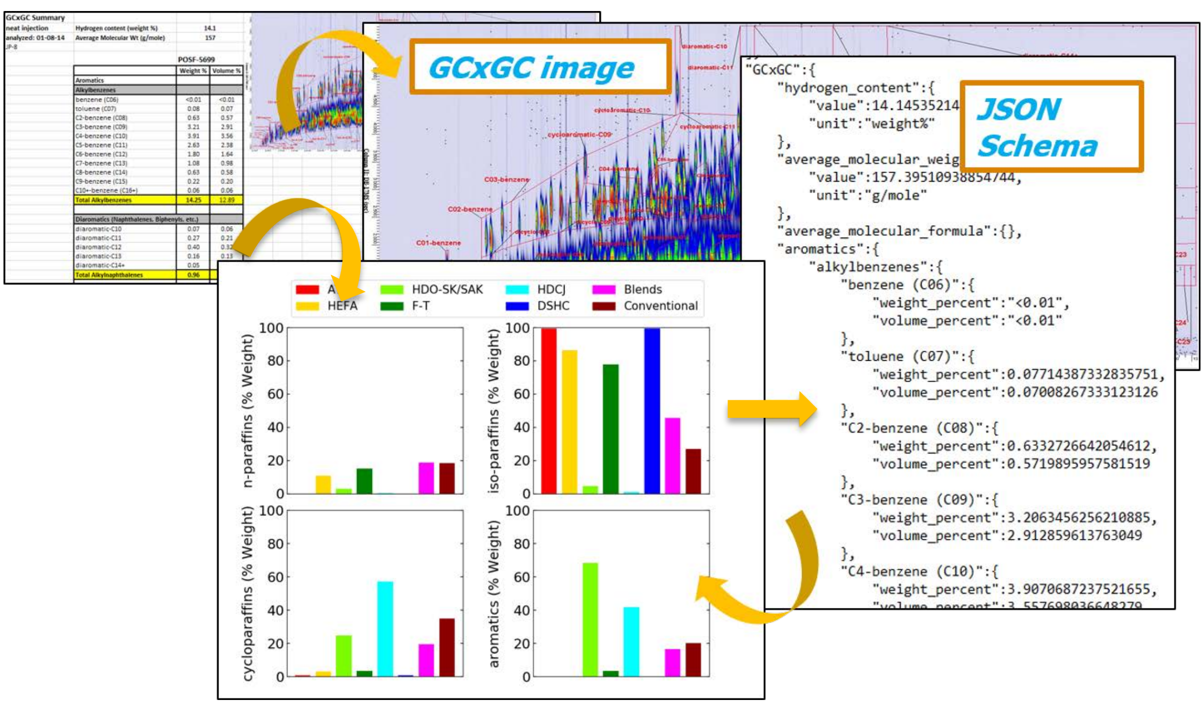


Figure 2. Conversion of GCxGC information to a dedicated JavaScript object notation (JSON) file using a Python script.

Currently, all of the data in the database have been converted into a JSON schema and can be analyzed in a flexible manner using Python scripts. An example of the analysis flexibility provided by the JSON schema (showing the fuel chemical composition) is presented in the lower right inset of Figure 2.

Generation II online database (summer 2019)

A beta version of the generation II database was launched online in the summer of 2019. The websites for both the generation I and generation II databases are shown in Figure 3. As data conversion to the JSON schema was nearing completion, we aimed to develop an online interface that could accommodate the new database format. The web version of the database has a three-tiered structure. On the outermost layer is an HTML-based web interface, as shown on the right side of Figure 3. All of the functionality of the previous database is maintained, and the security login features have been migrated from the previous version. The generation II database uses an interface that will also work with mobile and handheld devices. The inner layer houses the metadata, similar to the first database, and supports search functions for the user. The folder structure that was applied to organize the data in the first database was also retained in this version, allowing the user to access the data in this manner. The inner core contains the AWS database, which houses the JSON files, where the fuel data are stored. Duplicate data are saved on both the AWS and our department servers for security. File upload and comment functions for users have been carried over from the original version, and security oversight will be performed by the IT department at the University of Illinois.

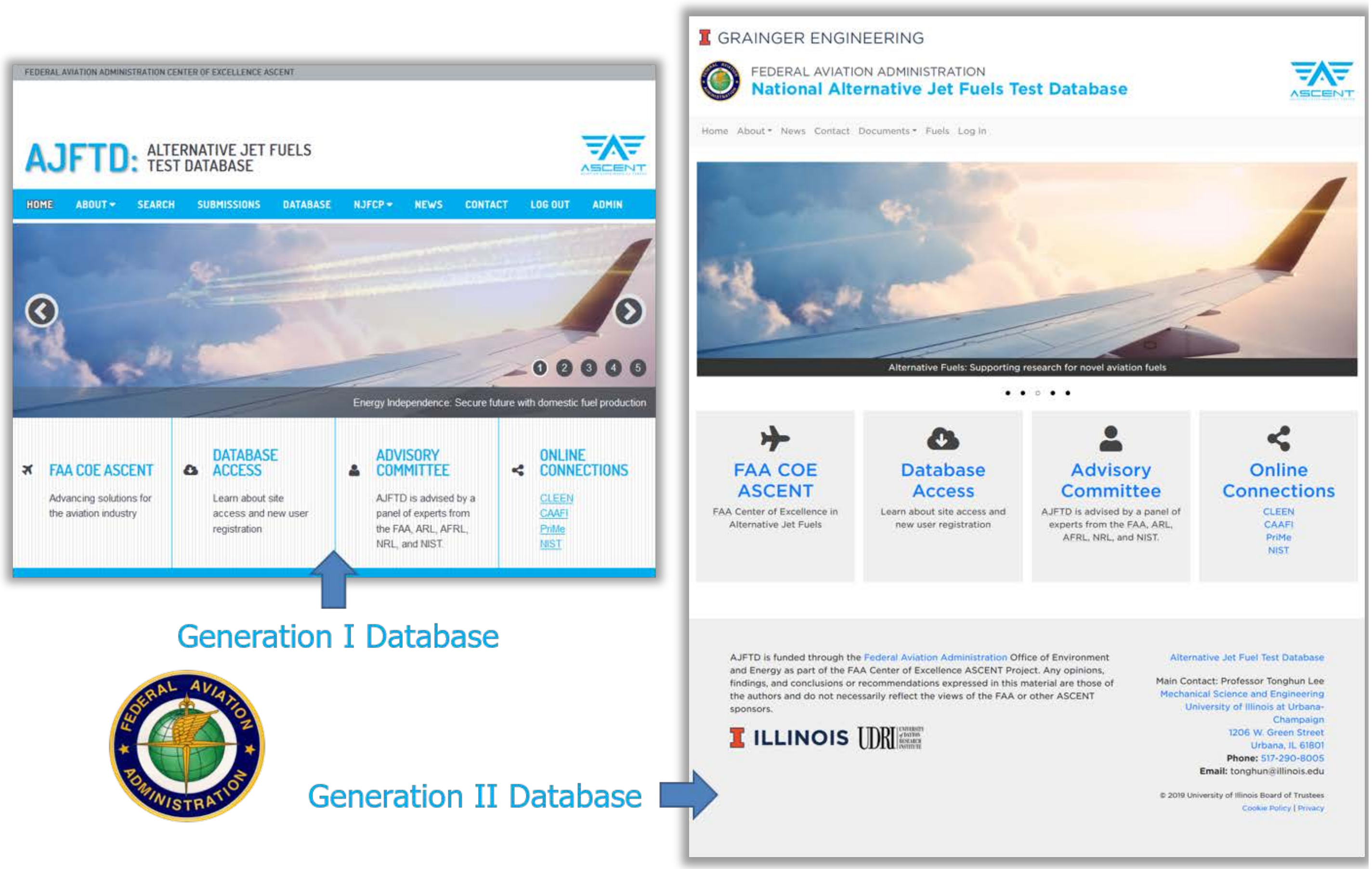


Figure 3. Generation I (left) and generation II (right) databases. URL: https://altjetfuels.illinois.edu/

As the entire dataset has now been converted to a nonrelational JSON format, we can apply a flexible range of analysis routines, as stated in the previous section. As a key aspect of the new database, we will provide a range of sophisticated analysis routines with which the user can directly interface on the website. The program for interfacing with the data files allows flexible analysis routines to be executed and easily modified. The first analysis routine to be built into the web interface will be a fuel comparison module. For example, a user can compare a specific property for a given fuel to the distribution for a subset of fuels in the database. An example of this routine is shown in Figure 4, where the flash point and kinematic viscosity of a specific fuel are plotted against those in the database. The user can also choose different subsets of fuels in the database to apply in the comparison. As shown in Figure 4, this analysis can be graphically presented through the website, and the user can download the actual data.

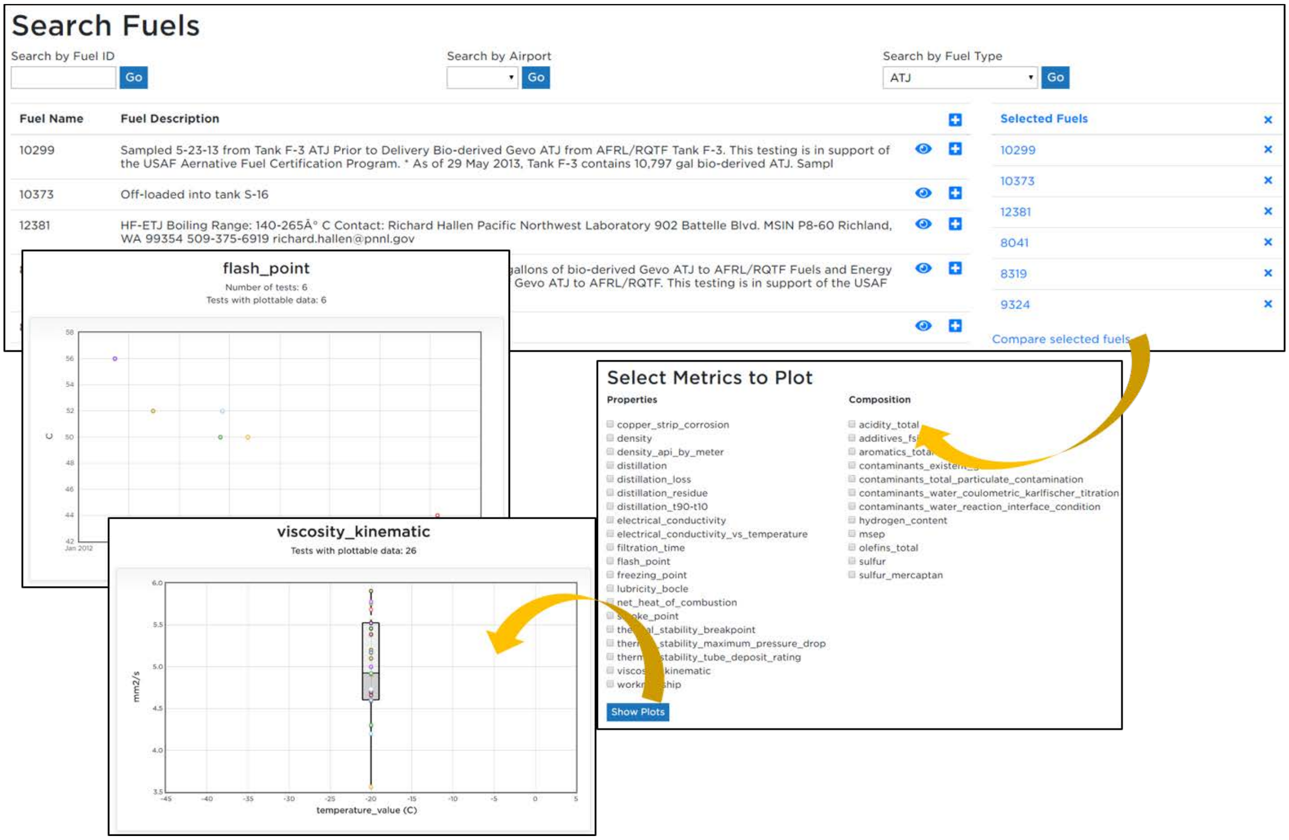


Figure 4. Interactive data analysis interface in the generation II online database.

In the future, we will expand the analysis capabilities of the database. We also envision applying more advanced correlation techniques based on machine learning to examine the correlations among different properties and their impact on the overall fuel performance. We have previously generated basic correlation factors for key properties using standard regression techniques. The application of machine learning may open the door to an extensive study of correlation factors that cover the entire property space of fuels in the database.

Milestone(s)

3 months

At the 3-month mark, we determined specific analysis routines for the database and decided to convert the data to a nonrelational data format in collaboration with the JETSCREEN program in Europe. The design of the generation II database was initiated, and the AWS database, which uses the DynamoDB infrastructure, was established at the University of Illinois. All of the Python scripts that were developed and used to facilitate the transition were documented.

6 months

At the 6-month mark, data conversion to the JSON format was fully completed and vetted. A preliminary analysis of the new nonrelational data was initiated, based on an examination of fuel property variations in our dataset. An internal beta test version of the database was completed, user testing commenced, and bugs in the web interface were identified. A technical and strategic roadmap for database integration with JETSCREEN was established.

9 months
At the 9-month mark, a fully working version of the generation II database was launched at https://altjetfuels.illinois.edu/. All data from the previous database were migrated onto the new system, and a secure online login was activated for all previous users. A vision of the joint JETSCREEN database integration was presented at the CRC meeting. The database analysis was extended to include Metron data from airports and GCxGC data.

12 months
At the 12-month mark, a preliminary online analysis tool was integrated with the database and made available to a general audience. This tool provides a graphical output of fuel property variations. Plans for expanding the online analysis capabilities and data download methods were drafted. A new data plotting and download function for users was added for fuel property comparisons.

Major Accomplishments

Conversion of data to JSON format (nonrelational digital format)
More than 25,000 fuel records have been converted to the JSON format, which can be easily accessed by most programming languages. GCxGC data have been obtained for >100 AFRL fuels and added to the JSON schema, and Metron data from airports (~14,000) have also been added. This step is the first toward establishing a foundational fuel database that will enable researchers to better analyze statistical correlations and compare property variations for alternative jet fuels. This format will also allow us to link the database with others, including JETSCREEN.

Launching of the generation II database website
We have launched a new version of the online database website at https://altjetfuels.illinois.edu/. The new website has enhanced capabilities, including (a) AWS DynamoDB integration using a nonrelational data structure, (b) online analysis tools for fuel property distribution visualization, (c) compatibility for small and handheld devices, and (d) flexible data upload and download capabilities.

Preliminary analysis
Using the new JSON schema, a fuel data analysis was performed to demonstrate the efficacy of the database for visualizing property distributions and identifying erroneous data and outliers. Composition graphs and distillation plots for alternative and conventional fuels were also produced as a starting point for the analysis. As more fuel data are obtained and added to the database, more meaningful and statistically significant analyses will be possible, including those based on machine learning methods.

Publications
N/A

Outreach Efforts
N/A

Awards
Anna Oldani (Graduate Student): DOT Student of the Year Award
Anna Oldani (Graduate Student): Society of Women in Engineering Award for Research Excellence

Student Involvement
This project was primarily conducted by one graduate student (Anna Oldani).

Plans for Next Period
In the next period, we intend to integrate online analysis tools based on 'big data' analysis. We will modify part of the database to be stored in the cloud for integration with the JETSCREEN program, which will provide a starting point for an international database. More details can be found in Task 2.

Task 2 – Preliminary Efforts to Link the Database with JETSCREEN

University of Illinois Urbana-Champaign

Objective(s)

The main objective of this task was to develop methods for integrating the current database with a similar system in the European JETSCREEN program. This effort was established through a series of discussions with the JETSCREEN team, where the benefits of such an endeavor were clearly identified. While the primary aim of this collaborative effort is to provide scientific data to support the approval process of alternative fuels, new directions, such as creating a foundational platform that can be used to integrate real-time fuel property data from flights, were also identified for future efforts. The major goals of this task are as follows:

- Identify benefits for integrating the database with JETSCREEN
- Converge on a single JSON schema for database integration
- Determine guidelines for comparing fuel properties for certification studies
- Determine the structure of a joint database and user interface protocols
- Determine specific near-term and long-term goals for database integration

Research Approach

Database integration with JETSCREEN

JETSCREEN is a European program spearheaded by the German Aerospace Center (DLR) and the University of Birmingham in the UK dedicated to providing fuel producers, air framers, and aeroengine and fuel system OEMs with knowledge and screening tools to streamline the alternative jet fuel approval process. The key objectives of JETSCREEN include (1) assessing the compatibility of fuels with respect to fuel and combustion systems, (2) quantifying the added value of alternative jet fuels, and (3) supporting research for optimizing fuel formulation to attain the full environmental potential of synthetic and conventional fuels.

Since the inception of the JETSCREEN program, we have communicated with the JETSCREEN leadership regarding approaches in which we can benefit from each other, as many of the goals of ASCENT align well with those of JETSCREEN. A timeline of milestones in our joint effort is presented in Figure 5. As the most important aspect of this specific project, JETSCREEN is establishing a database much like the system discussed herein. In the later stages of 2017, we concluded that it would be mutually beneficial to integrate parts of the two databases into one system and to establish a foundation that could, in the future, encompass additional database links around the globe. In December 2018, at the annual JETSCREEN meeting, we presented specific goals, both technical and strategic, regarding this integration process. From a technical perspective, we put forth a common nonrelational JSON schema and other details regarding the database structure, interface, analysis routines, etc. It was also determined that while the data location would be common for our database and JETSCREEN, the interface and analysis routines would be different and customized to suit the needs of either party. Both our database and JETSCREEN would develop separate online portals for accessing the data. From a strategic perspective, we plan to gear the database towards two critical goals: (1) fuel certification analysis and (2) inclusion of fuel properties from airports (described in more detail below). These goals were presented at both the CRC and IASH meetings in 2019.

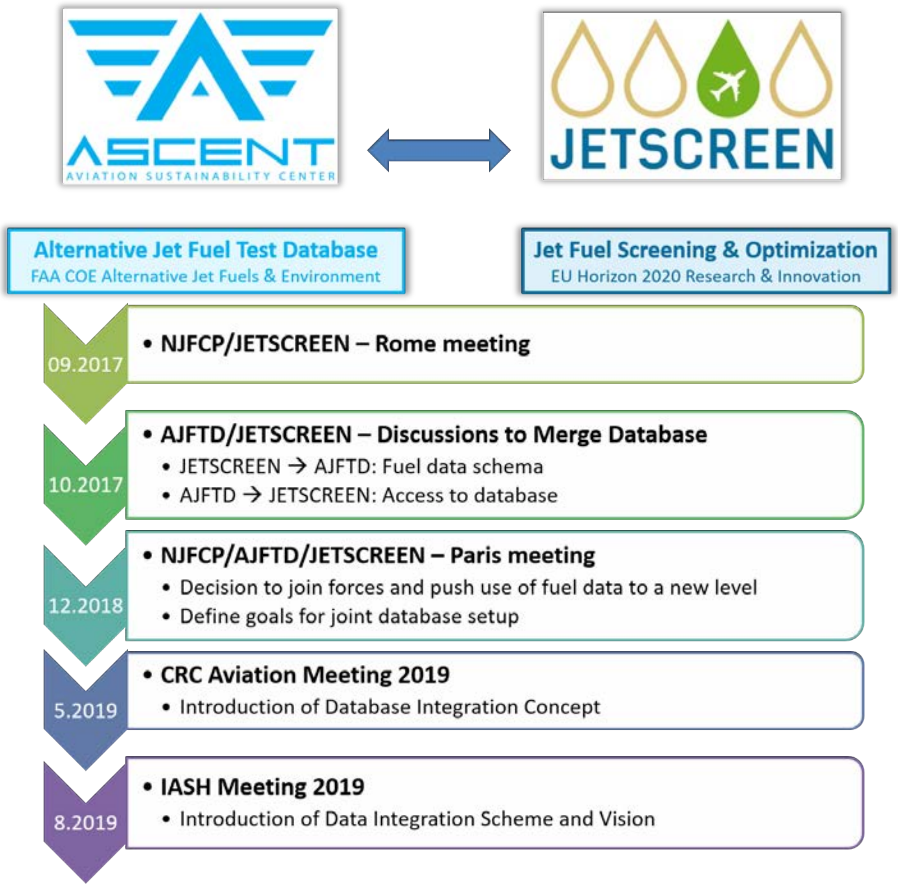


Figure 5. Timeline of milestones relevant to the integration of our database with JETSCREEN.

Fuel property analysis to support certification

This database makes a critical contribution by providing a measure for comparing new fuels to current fuels in the database. A combination of whisker plots and data points was chosen for optimal visualization of this comparison, as illustrated in Figure 6. Here, properties from a specific new fuel blend are compared with those for a subset of fuels in the database. The center point of the box shows the median, and the interquartile ranges are shown as the edges of the box (left and right). The solid whiskers show the distribution of fuels from the database while the dotted whiskers show the limits of the specification. The green data points designate properties that are within the range of fuel properties in the dataset, while the yellow points show properties that are outside the range of the database but still within the specification guidelines. The red points fall outside of the specification entirely. Both our database (generation II) and the JETSCREEN online interface will utilize this comparison method, where users can choose to compare a selected fuel with our entire database or a subset of fuels.

Fuel specification	Min	Max	Fuel Sample
Colour			30
Acidity (mg KOH/g)	0	0.015	0.002
Aromatics IP 156 (%vol)	0	25	1.8
Sulphur (%mass)	0	0.3	0.018
Mercaptan (%mass)	0	0.003	0
IBP (degC)			148.8
10% (degC)		205	169.8
50% (degC)			198.7
90% (degC)			235.1
FBP (degC)		300	251.9
Flash point (degC)	38		41.5
Density @15degC (kg/m3)	775	840	759.6
Freezing point (degC)		-47	-59
Viscosity @-20degC (cSt)	0	8	3.885
Smoke point (mm)	5		50
Naphthalenes (%vol) if SP > 25mm		3	0
Specific Energy (MJ/kg)	42.8		44.023
Existent Gum (mg/100 ml)		7	1
MSEP	85		99

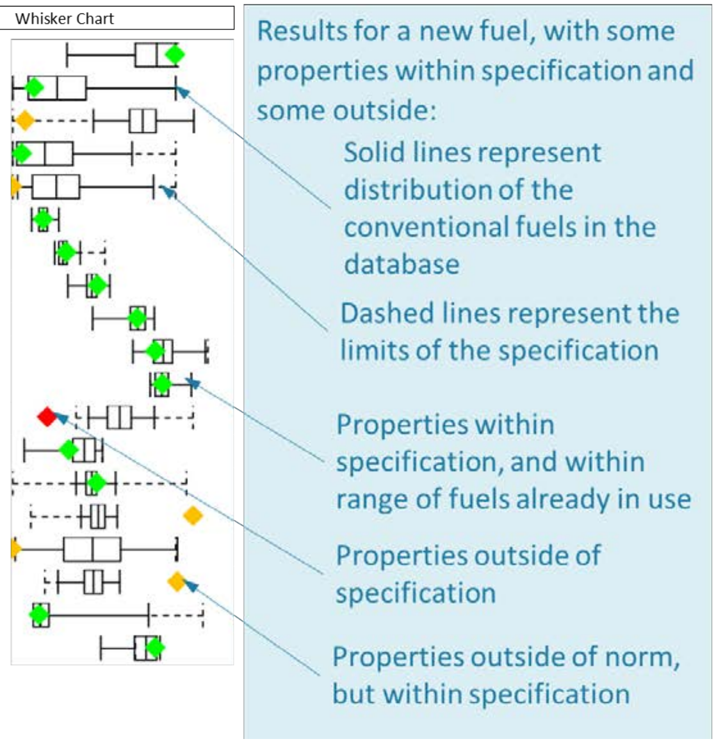


Figure 6. Comparison of fuel properties with those in the database and ASTM specifications. The red data point indicates a value that is outside the specifications.

In addition to the fuel property analysis shown in Figure 6, the nonrelational schema utilized for the database allows us to deploy a range of advanced analysis methods, including machine learning algorithms, that can seek different correlation factors between properties and fuel performance. Even a general correlation between properties may reduce the tests required for the approval process, and an analysis of GCxGC data could shed light on the relationship between fuel performance and fuel blend chemical composition. To this end, JETSCREEN and our team will launch a machine learning exercise to identify critical correlations between properties that have not yet been determined through previous methods.

Figure 7. Impact of integrating real-time data from airports regarding fuels used in actual flights.

Integration of actual fuels used at airports

A new direction was identified during the early stages of discussion with JETSCREEN: to seek a method for compiling data on fuels utilized in commercial flights. This direction was motivated by two aspects. First, in the process of compiling the database, we came across airport data collected by Metron Aviation through an earlier ASCENT project. It became apparent that this type of data was very valuable, providing a roadmap for tracking real fuels being deployed in the field. We were surprised to find that as the data were being generated, no standard archiving method was applied, and the data were being discarded after quality control of the fuel had been performed. Second, as we identified the long-term impact of this database (sample shown in Figure 7), the need for data on fuel that was actually being used, both in real time and compiled over the long term, became increasingly evident. Such data would not only provide a means for assessing the quality and variation of fuels currently in use but, over time, would also provide an archive of changes in overall chemical compositions and physical properties. This type of data can greatly aid in the analysis of the environmental impact of fuels over time.

Milestone(s)

Throughout this year, we have worked with the JETSCREEN team to develop a plan for merging a portion of the database. We envision that this step will establish the foundation for an international platform, where various nations can both upload fuel data and utilize data from other nations. We have set forth several motivations and strategies to complete this task and have established firm guidelines for use in this collaboration. We plan to present these plans and acquire reviews from the academic community in the next year and will then make the necessary modifications.

Major Accomplishments

JETSCREEN integration vision

We have established a technical and strategic vision to integrate our database with the JETSCREEN database. The new database will not only provide support for the integration of alternative fuel blends, but will also provide a foundation that can integrate data for fuels used in commercial flights. In this manner, the system will amass data that can optimize both operational efficiency and safety in the future.

Publications

Oldani, Anna. (2019). Alternative jet fuel variation and certification considerations. In progress.

Outreach Efforts
N/A

Awards
Anna Oldani (Graduate Student): DOT Student of the Year Award
Anna Oldani (Graduate Student): Society of Women in Engineering Award for Research Excellence

Student Involvement
This project was primarily conducted by one graduate student (Anna Oldani). Oldani surveyed the data, interacted with the data sources, and created strategies to integrate the data into the database. Oldani also developed a web-based portal for implementation of the web interface and analyzed the available data to evaluate the property variance. Oldani presented these results at conferences and FAA ASCENT meetings. Anna Oldani completed her PhD in April 2019 and began working at the FAA Office of Environment and Energy in September 2019.

Plans for Next Period
Database integration and capture of real-time fuel use data
In year VI of the database project, we will implement fully extended data evaluation activities using the new nonrelational structure and will integrate the database with JETSCREEN. Additionally, we will engage with airports to evaluate the potential for integrating real-time data. New advanced analysis methods will be made available online.

- Evaluation and analysis of all data, including GCxGC, Metron, and PQIS data
- Utilization of machine learning algorithms for data analysis (property correlations)
- Integration of a common database with JETSCREEN, with separate user interfaces
- Interaction with airports to assess pathways for capturing data on fuels in use
- Inclusion of additional data and continued integration of fuel data from all sources