Check for updates

Advanced Fuel Property Data Platform: Overview and Potential Applications

Simon Blakey[1,2]*, Bastian Rauch[3], Anna Oldani[4,5] and Tonghun Lee[4]

[1]Department of Mechanical Engineering, University of Birmingham, Birmingham, United Kingdom, [2]Department of Mechanical Engineering, University of Sheffield, Sheffield, United Kingdom, [3]DLR, German Aerospace Center, Institute of Combustion Technology, Cologne, Germany, [4]Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Champaign, IL, United States, [5]Current affiliation, Federal Aviation Administration, Washington, D.C., United States

This report outlines the establishment of distributed databases for management and integration of current and future aviation fuels. Aviation fuel property and performance data has been gathered for many years in public and company specific fuel surveys. These surveys are suitable for use as overall quality control information and for monitoring changes and trends in the fuels in used for flight. In recent years, significant data has been generated for alternative fuels as part of the due diligence of their approval for use through ASTM D4054, including those outside of the specification. Recently, this data, along with fundamental chemistry data has led to the creation of the Fast Track route for fuels approval when the fuel is constrained to a necessarily narrowly defined composition. The data behind these developments are often stored in a disparate, unindexed way, resulting in their underutilisation for a range of research, engineering design, specification, and in service quality control applications. To make the best use of this data, we present a scalable, Json based format for the storing of fuels data. This concept has been proposed by the Horizon 2020 Jet Fuel SCREENing and Optimization (JETSCREEN) project in conjunction with the Center of Excellence for Alternative Jet Fuels and Environment (ASCENT) programme. We have worked collaboratively to develop a joint database which currently contains data from around 30,000 conventional and 400 alternative fuels/fuel blends from a range of European and United States of America (U.S.) lead research programmes and data sources. This database can be used for a variety of purposes, both in conjunction with, or in isolation of commercially sensitive data with a greater degree of restriction. We present a number of test cases for how we see this model for data storage could be used for the benefit of all. We invite further suggestions as to how this approach could be used and welcome opportunities to work with the wider fuels community to develop this idea further.

Keywords: aviation fuel, alternative fuels, fuels approval, database, fuel properties

INTRODUCTION

The civil aviation sector has spent much of the last 60 years optimizing the design of aircraft and engines to reduce fuel consumption, lower CO_2, noise and NO_x emissions. The sector has had significant success in this undertaking, however during this period the composition and properties of aviation fuel have been treated as a bought in commodity, and have essentially remained unchanged. This has meant that the sector has been optimizing changes to the fuel systems in aircraft and ground handling hardware around a range of average, or for specific properties, the worst case fuels available.

These conventional fuels, coming from sources identified in **Section 4** of the DEFSTAN 91-091 (Ministry of Defence 2016) and Section 6.11 of the ASTM D 1655 (ASTM International 2020) are: "Aviation turbine fuel, except as otherwise specified in this specification, shall consist predominantly of refined hydrocarbons derived from conventional sources including crude oil, natural gas liquid condensates, heavy oil, shale oil, and oil sands." These fuel types are seen as those suitable for gas turbine powered aviation and as long as these sources produced fuel which met the specification, they were permitted to be used, regardless of the actual hydrocarbon chemical composition and performance properties of the fuel. In addition to the hydrocarbons, the specification places requirements on the cleanliness of the fuels and the concentrations of hetroatomic species, metal and water content. These fuel sources have not significantly changed since the beginnings of the jet age, and as such is the "Jet A-1 everybody knows," based on accumulated experience. This means that much of the risk in the use of these fuels is mitigated through the use of this experience and trusted, standardized specification measurements used in the above standards, and the standards represent a batch certification of the fuel as safe for flight.

Throughout this period, fuel user groups have carried out surveys of fuel quality based around the available fuel property data as part of the Jet A-1 specifications, and this data has been reported in regional and global fuels surveys. Key examples are those carried out by the United Kingdom MoD (later the Energy Institute) (Energy Institute and QinetiQ 2014) and the Petroleum Quality Information Service (PQIS) annual reports (The Defense Energy Suppor 2009), amongst others, which report the variability of fuel in use currently. This data is limited, in the extent that it reports the specification performance of the fuel to the above standards without detailed information of the chemical composition of the fuel. As the PQIS survey includes a wider range of fuels for applications outside of aviation, it also contains additional data such as H/C ratios and Cetane index results. These survey reports are supported by the beyond specification information provided by the CRC world fuels survey (Hadaller and Johnson, 2006). And the CRC Aviation Fuel Handbook (Coordinating Research Council 1983), which is a valuable and commonly used reference within the industry, despite it representing Jet A-1 fuel by a single line, as shown as shown in **Figure 1** along with the range of fuels within fuel surveys (Coordinating Research Council 2014). These combined data are invaluable in assessing what fuel is being flown on at the present time and also in keeping track of the longer term trends in fuel quality over time. The information captured in these reports is of great importance to the sector, however, it is often contained in paper based reports, or the electronic equivalent (such as pdf files), which limits its usefulness to the sector as a tool to assess the "fit" of any alternative fuel, or indeed any conventional fuel from a novel source.
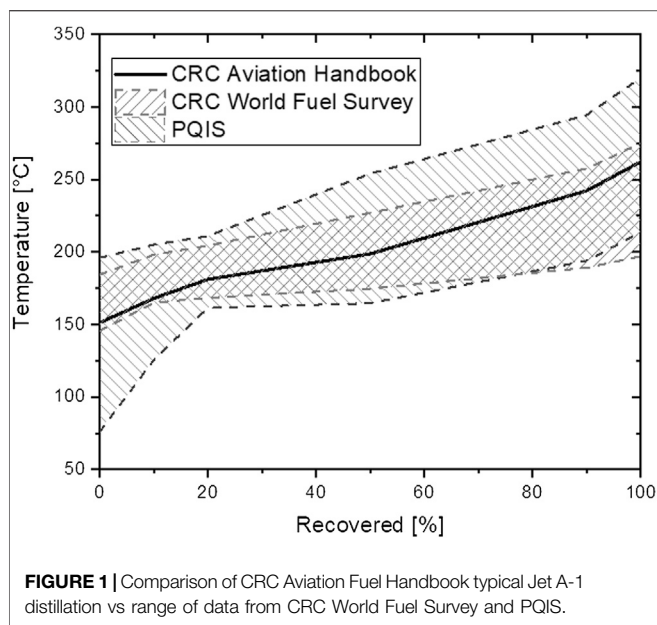
During a brief period at the beginning of the 1980's and a more sustained manner since the early 2000's an increasing range of alternative fuels have been proposed which, following thorough testing, have been demonstrated to be technically suitable for use in civil aviation, and approved for use through the ASTM D4054 process (ASTM 2020). These fuels include Sustainable Aviation Fuels (SAFs) which are produced from sustainable sources. Importantly the sustainability criteria for SAF are not assessed in the D4054 process, which is exclusively a technical suitability assessment. These fuels have been assessed in a far more technically rigorous way than those fuels derived from fossil sources highlighted above. The D4054 is a robust process, and one which follows a strict management of change in the expansion of fuel sources away from the "Jet A-1 everybody knows." Unfortunately, at present, much of this technical information is stored in reports and files in a way which makes the data very hard to access and use. As such there is a risk that this valuable information is not being used to accelerate the screening of any future candidate fuel, or in the assessment of engine and airframe performance with modified fuel properties.

METHODOLOGY

Common Philosophy

The Horizon 2020 Jet Fuel SCREENing and Optimization (JETSCREEN) (Rauch 2020) and Center of Excellence for Alternative Jet Fuels and Environment (ASCENT) (altjetfuels 2015) projects had both independently proposed methods for increasing the usefulness of this information through a common data schema, an online source for data respectively. The projects identified an opportunity through the publically funded work to share data. As part of this opportunity, it was necessary to develop a common philosophy for the storage and use of data. This paper details this common philosophy and goes on to present some examples of usage in the hope that future discussions develop these ideas further.

The overarching concept of the JETSCREEN database project is to make available the public data generated in previous research activities and in the current JETSCREEN project, in a form that is human readable and can therefore be used to produce statistical and comparative analysis of any candidate fuel. It is envisaged that this assessment can streamline some of the early screening processes of the D4504 process for fuel approval, as captured by the "Tier Zero" or "Tier Alpha" concepts presented by JETSCREEN and the FAA (Heyne et al., 2021). This dataset will start with the chemical hydrocarbon composition of the fuel, using a method such as GCxGC to identify the molecular families of molecules present in the fuel, the specification properties and fit for purpose data required for approval, but will go onto include

FIGURE 1 | Comparison of CRC Aviation Fuel Handbook typical Jet A-1 distillation vs range of data from CRC World Fuel Survey and PQIS.

more diverse data on a fuel's behavior during the course of the JETSCREEN project. Importantly, where data already exists in the public domain, a candidate fuel can be compared to the fuels already in use, those approved, and importantly those not approved and help to develop structure-property relationships for molecules in jet fuel.

Where possible, the database will provide information which can be used to develop and validate statistical, empirical and fundamental models linking the chemical and compositional details of a fuel to its specification and performance properties. Importantly, it should also be possible to use the models developed through this process to predict compositional information based on desired performance properties. This is an active area of research and many sources of data and analysis studies have been made using available fuel property data (Dryer et al., 2014; Moses 2017; Heyne et al., 2019; Heyne et al., 2022) amongst many others.

As such the database proposed can provide a single knowledge base of fuels which are in use today, have been through the approvals process and those from the research community where data is available. This approach will provide significant gearing for the use of this data in the assessment and screening of candidate fuels in the future particularly if the users can determine the quality control on the uploaded data, an oversight role that the authors currently perform for their respective databases. In the preparation of this work, it became clear that the database could have significant uses beyond fuel pre-screening and should be of interest to the wider aviation fuels community as a resource for conducting any fuel related monitoring and development studies. This schema provides a structure for open science and the sharing of data which encourages advancement and the rapid adoption of new technologies in the field of fuel properties by promotes diverse, just and sustainable outcomes for all stakeholders (Grahe et al., 2020).

Data Schema

The sharing of data is limited to that data which is available in the public domain. Clearly, many datasets are private and not available for sharing, however if they share a common file storage schema or format, the results of these datasets may be rapidly integrated to provide the user a more statistically significant set of fuels data within individual organizations. The current joint database between JETSCREEN and FAA is based on such a common storage schema, where non-proprietary data is shared through a common cloud server as shown in **Figure 2**. Both JETSCREEN and FAA can access this server automatically to upload and download data through an hourly or daily sync process.

Shared data is stored using a mongo DB database structure using a common standard JavaScript Object Notation (.json) format. This formatting is beneficial as it is an unstructured method for storing data and can incorporate a high degree of flexibility whilst providing a standard, human readable format which can also be easily interpreted by computer code. A live schema for fuel data storage is maintained by the JETSCREEN consortium at the following URL: https://github.com/JETSCREEN-h2020/FuelDatabase/wiki/JETSCREEN-Schema-philosophy.

The schema of a database is the organization of data i.e. how a database or the data exchange file is constructed. At this URL, the organization of fuel data in the form of a JSON file is described. The structure presented here is version 2.0 of the data schema.

The data schema is a living document and will grow and mature with the projects and its use. As such, the current documentation will not be listed in this paper, however, links to a github site which provides access to the latest version will be embedded into the document. An example of the current schema is shown in **Figure 3**.

The basic schema of fuel data is divided into three parts:

(1) A Header section: with metadata about the fuel and authors. (https://github.com/JETSCREEN-h2020/FuelDatabase/wiki/Header).
(2) Composition section: example aromatics, contaminants etc (https://github.com/JETSCREEN-h2020/FuelDatabase/wiki/Composition).
(3) Properties section: like acidity, distillation, flash point, etc (https://github.com/JETSCREEN-h2020/FuelDatabase/wiki/Properties).
  • Use issues
(1) Conventions.
  - https://github.com/JETSCREEN-h2020/FuelDatabase/wiki/Convention
(2) Data quality.
  - https://github.com/JETSCREEN-h2020/FuelDatabase/wiki/Data-Quality

Data Sources

The principle sources of published data used in this dataset are a combination of survey results of conventional fuels in service today, research reports and publications on alternative fuels. The major drawback of the fuels survey results is that they contain
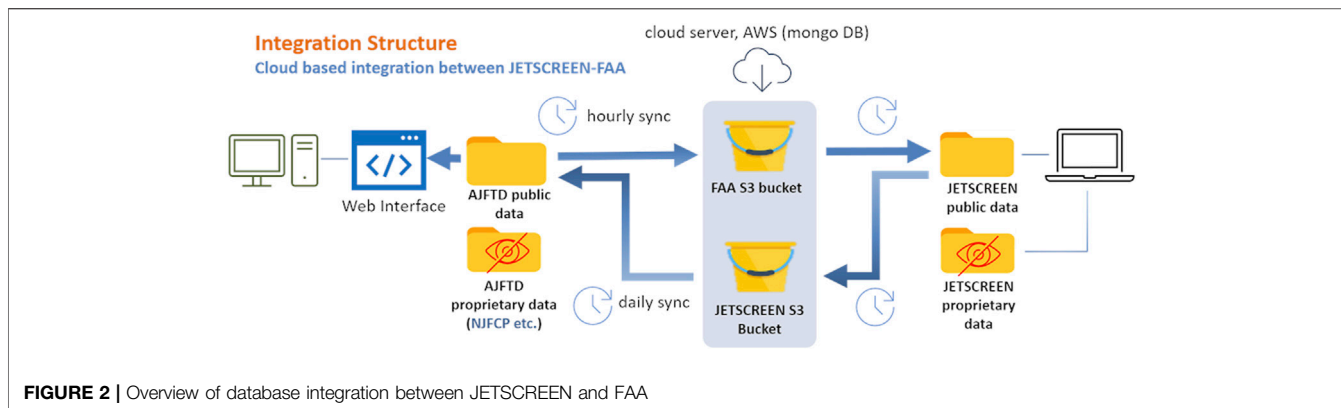
FIGURE 2 | Overview of database integration between JETSCREEN and FAA

FIGURE 3 | Excerpt of the standardised fuel property schema showing the formatted structure of the data recorded including fuel meta data.

specification are represented in the database. The fuel data represents a variety of manufacturers, and all data was acquired from five main sources: Metron Aviation, the National Jet Fuel Combustion Program (NJFCP), Air Force Research Laboratory (AFRL), Naval Air Systems Command (NAVAIR), and the European program JETSCREEN.

Table 1 below outlines a sampling of fuels from AFRL, available on the AJFTD website. This table demonstrates the breadth of fuel types available to users, from a wide range of fuel types to a variety of fuel manufacturers. The selected property and composition categories included in the table are included to exemplify the potential range and variability of fuel properties and compositions observed between different fuels, variability which in some cases spans a large range of acceptable property or composition limits as outlined by ASTM D7566 standards. This kind of variability underlines the usefulness of a centralized, extensive database in elucidating the degree of the aviation fuel industry's heterogeneity, especially considering how ongoing development of new certification pathways will inevitably increase diversity in this arena.

Fuel data is often difficult to obtain from manufacturers due to policies protecting proprietary information. This leads to data gaps in the database. Fuel samples from the most recent Annexes to D7566 are not as well represented as older ones. Additionally, the acquisition of fuel data from a variety of sources inevitably leads to data sparsity and inconsistencies in data categories among different fuel samples. Critical next steps for the online database include ongoing database integration with international programs, interception of fuel samples from domestic airport supply chains, re-organization of the database structure to optimize navigation capability and interactive features, and the incorporation of data analysis tools like machine learning algorithms. These improvements will better equip the website for user adoption.

The well-established example of the risk associated with using average data rather than individual fuel data is the calculation of dynamic viscosity from the average density and kinematic viscosity of a fuel data set and the average dynamic viscosity from the individual densities and kinematic viscosities of all fuels in the dataset, as shown in **Table 2**. Although the average dynamic viscosities calculated by two methods are similar, the value calculated from the individual densities and kinematic viscosities also contains standard deviation

little or no information regarding the chemical composition of the fuel. There are many thousands of unique fuels across these published data sources which can be integrated into a common database and, their inclusion gives the database and any user an understanding of in service fuel quality. This is significantly more valuable than an understanding based on a single baseline fuel and its associated performance properties. Importantly, this data should be stored on file as specific properties of a single fuel, and not as statistical averages of fuel properties.

The AJFTD contains a vast library of information relevant to the alternative jet fuel industry and is accessible via https://www.altjetfuels.illinois.edu/. While its main feature is its library of over 25,000 different samples of domestic and internationally sourced fuels, other data categories available include documents on chemical kinetics mechanisms, aviation emissions, relevant publications and literature, and experimental testing results. Alternative fuels from all Annexes of the ASTM D7566

TABLE 1 | Examples fuels available on AJFTD website.

Fuel Type	POSF/AJFTD Name	Manufacturer	Aromatic Content [%vol]	Density [kg/m^3]
JP8	4751	N/A[a]	19.2	804
Jet A	10325	Shell	17.4	803
FT	5642	Sasol	0.7	762
FT	7629	Sasol	1.9	760
HEFA	5480	Syntroleum	0.6	762
SIP BLEND	50% SIP	N/A[a]	8.4	793
ATJ-SKA	ATJ-SKA	N/A	19.4	786
ATJ	7695	Gevo	0.0	760
ATJ BLEND	7700	Gevo	9.2	782
CH-SK	CH-Kerosene	N/A[a]	19.7	805
HC-HEFA	13784	IHI	0.0	782

[a]*Entries marked "N/A" denote information not available on the database.*

TABLE 2 | Example of error introduced by calculations based on average values from property databases.

	Calculated from individual densities and kinematic viscosities of all fuels in the dataset (Ns/m^2)	Calculated from average density and kinematic viscosity (Ns/m^2)
Minimum	0.001011	n/a
Maximum	0.002353	n/a
Standard Deviation	0.000236	n/a
Average Dynamic Viscosity	0.001426	0.001425

information which is missing in the conventional approach. This greater detail is important for a number of stakeholders and like **Figure 1**, reinforces the view that it is misleading to using average values for fuel properties. Engineers using these properties for their work should take into account the statistical information such as standard deviation calculated from a database of these properties.

The growing volume of data on the performance of alternative fuels provided through the research literature, D4054 reports and aviation fuels research projects such as JETSCREEN and the U.S. National Jet Fuel Combustion Program (NJFCP) (Colket and Heyne 2021) are more useful for the stated goals of the database, as they contain detailed compositional information as well as specification and performance data. However such data is much more sparse. Current estimations are that there are around 150–180 fuels with sufficient data from these programmes to be included at the start of this exercise. Included within these studies are a number of conventional fuels for which this additional analysis has been conducted, including GCxGC data and fit for purpose testing fuel specification properties, as well as fundamental combustion properties (Oßwald et al., 2021). This smaller sub-set of research fuels is very useful in bridging the gap between the larger, more statistically useful conventional fuel surveys and the more detailed datasets from research into fuel performance.

It is hoped that this schema can be adopted in future programmes and surveys so that fuel composition, specification and performance data can be presented in a truly interoperable way to integrate into existing and future tools and workflows. This level of adoption would greatly increase the usefulness of any fuel property data generated. In the following sections, several usage cases for this database are presented. It is fully envisaged that these usage cases can be expanded upon significantly by individual users with particular fuel data requirements.

APPLICATION

Screening and Safety

In line with the original objectives of JETSCREEN, the fuels stored in the fuels database can be used to compare the specification and performance properties of a new candidate fuel with the fuels present in the database, offering a rapid comparison with existing fuels and presenting the data in a graphically simple representation, easily understandable by the user. This is a method to assist the acceleration through the early stages of pre-screening and assist fuel producers with access to fuel property data [xi]. **Figure 4** shows a comparison of a 100% HEFA fuel from the E.U. project 308,807: Initiative Towards sustAinable Kerosene for Aviation (ITAKA), as represented by the individual data points for each property, the fuels used in the United Kingdom in 2014 (as represented by the whisker plots, showing the minimum, 1st quartile, mean, 3rd quartile and maximum value for each property) and the conventional fuel specification, in this case DEFSTAN 91-091 (dashed lines). It is clear that the candidate fuel would not comply with this specification as it is above the 50% blend limit, and is shown here for example as some of the properties of the fuel do fit within the specification without blending. This clearly shows the areas where the candidate fuel is within the specification (indicated by a green colour), with the
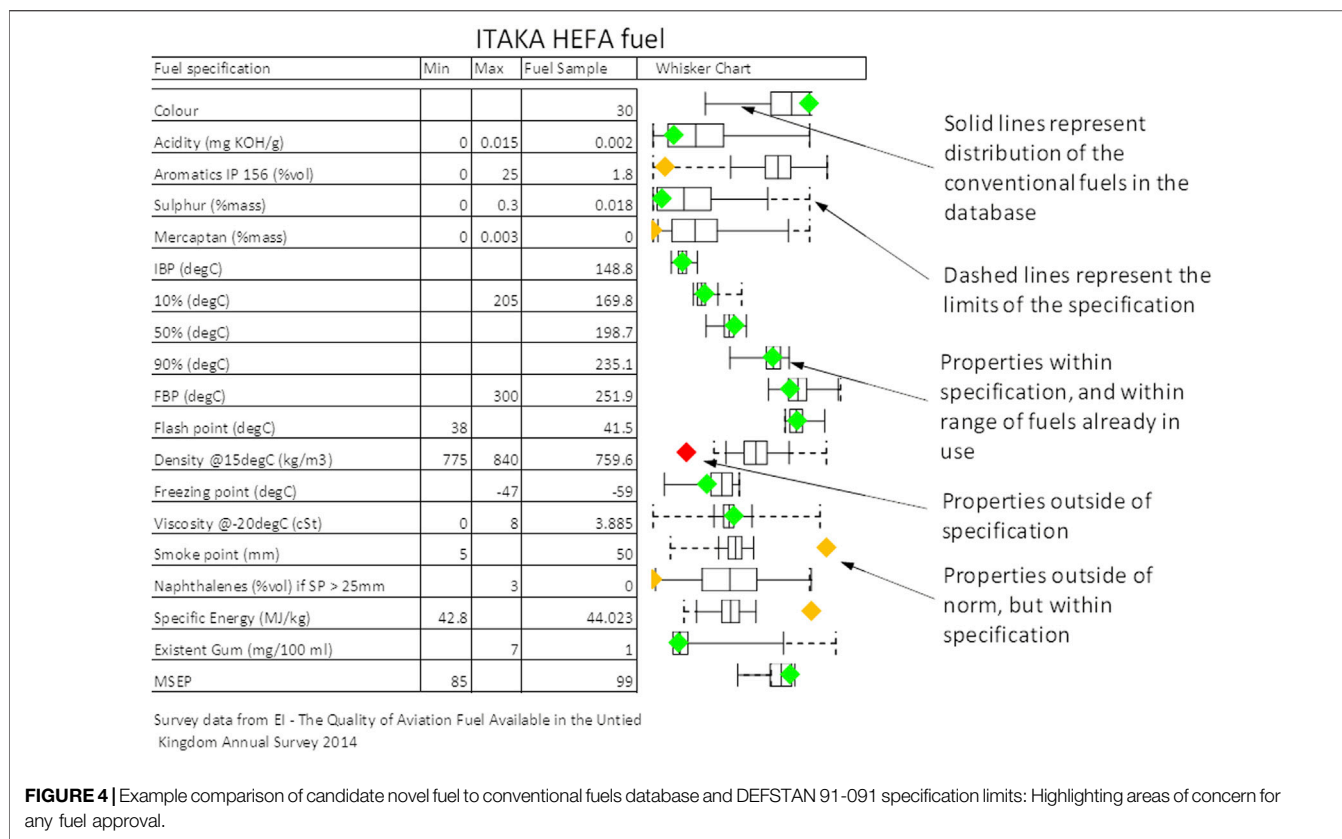
FIGURE 4 | Example comparison of candidate novel fuel to conventional fuels database and DEFSTAN 91-091 specification limits: Highlighting areas of concern for any fuel approval.
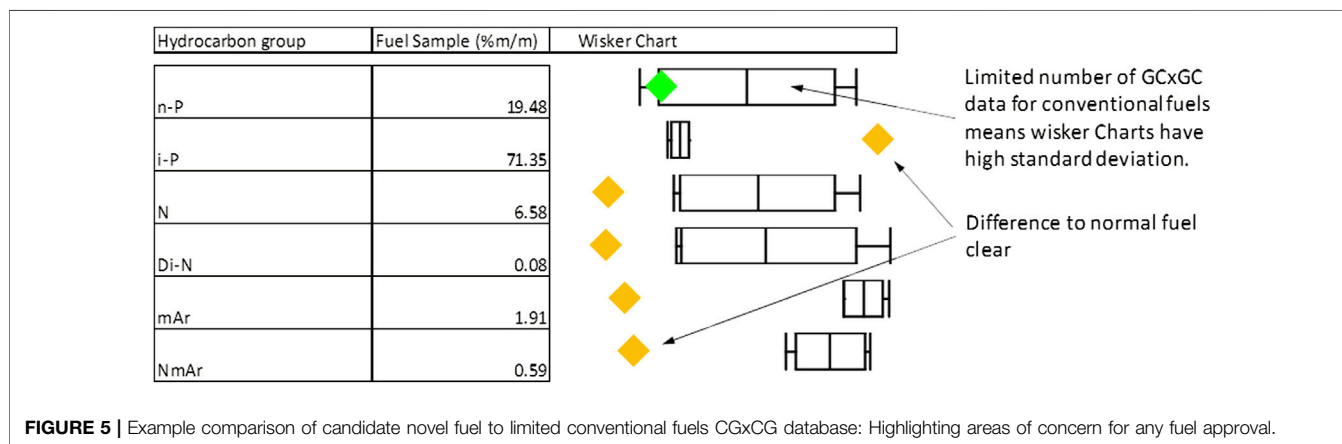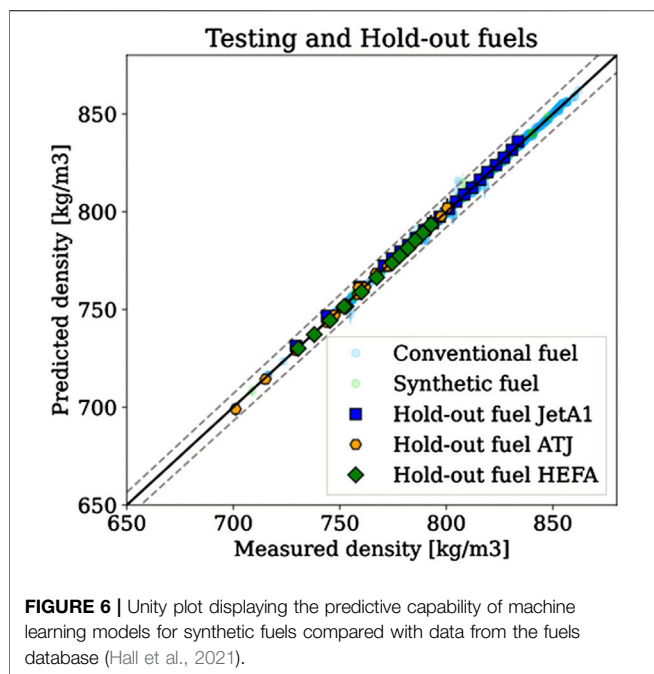

FIGURE 5 | Example comparison of candidate novel fuel to limited conventional fuels CGxCG database: Highlighting areas of concern for any fuel approval.

specification but outside the norm for conventional fuels (indicated in orange) and where the specific property is outside the specification, (indicated in red). For this candidate fuel, it is immediately obvious that the only real spec failing would be the density of the fuel. This is due to the low levels of aromatics present in the ITAKA product. Although the other specification properties are still within the specification, they are well outside of the norm–specifically, low aromatics, low sulfur, high smoke point and high calorific value. In many respects, this understanding is implicit for the fuels experts of the aviation sector. However, as the number of fuel producers

increases with the various feedstocks proposed to produce aviation fuel from unconventional sources, it is necessary to communicate the particular requirements of the aviation specifications to an increasingly wider audience. For such information sharing, simply understood graphics such as the whisker comparison plots, and a traffic light colour scheme are essential.

Also **Figure 5** shows a comparison of the GCxGC composition of the fuel compared with the limited range of conventional fuels already in the database. This comparison also shows clearly where the ITAKA fuel is outside of the norm. Importantly, as there are

FIGURE 6 | Unity plot displaying the predictive capability of machine learning models for synthetic fuels compared with data from the fuels database (Hall et al., 2021).

no specification limits for a GCxGC composition, therefore there are no returned red data points.

Both **Figure 4** and **Figure 5** can be used to consider the maximum acceptable blending ratio of SAF with conventional fuels, indicating that for most properties, there would be a large variation in acceptable blend depending on the conventional blendstock specification results DEFSTAN 91-091 for specification limits.

Engineering and Science

It is envisaged that as the uptake of SAF increases, the usefulness of such fuels data will increase as there will be a slow drift in fuel specification properties over time as greater volumes of SAF are blended with conventional fuels. In order for this process to be properly monitored and controlled, availability of specific fuel by fuel data will be required in order to exploit the largest benefit from the SAF.

In addition, a common fuel schema and database could allow a quicker identification of specific fuel properties if further investigations are required into following a fuel related system or component failure in the supply system or on-board.

Access to larger datasets of specification and further fit for purpose properties of fuels in a common format would greatly enhance the ability of engineers and scientists to perform analysis of fuel behaviour and performance, increasing the understanding of the link between fuel composition and fuel performance.

This would also facilitate the development and validation of statistical and more fundamental models of fuel properties adding to the chemo-informatics tools which are being developed to assist in the early screening of candidate fuels for the approvals process. This can also be used to enhance the development and production of alternative fuels, particularly in using these

developed tools to point towards optimum fuel compositions for performance in flight.

The statistical analysis and feature detection for aircraft related fuel properties can facilitate the design of aircraft components impacted by the fuel performance. As optimisation of the engine and airframe continue further, the fuel systems are likely to become increasingly sensitive to changes in fuel composition. Therefore access to fuels data which are easily integrated into design tools would improve the workflows of the design process.

This dataset is already in use within the JETSCREEN project to develop and validate machine-learning and other tools to predict important fuel specifications and performance characteristics as shown by the example in **Figure 6**, showing the prediction of fuel density from the GCxGC compositional results alone for a wide range of different aviation fuels (Hall et al., 2021). The term hold-out is used to indicate data used for assessing the machine learning model after the training step has been completed. Furthermore, as considerable amount of data is available for conventional and synthetic fuels, the predictive capability (accuracy, prediction uncertainty and model reliability) of models can be assess systematically over the potential application domain.

Figure 7 shows a comparison of a range of fuels from the ASCENT database, looking for correlations between fuel compositions and emissions performance. This shows the suggestion of correlations between Smoke Point and DCN, but more significantly between Smoke Point and H/C ratio. Such correlations can be made for any select group of fuels, and efforts are being made to adapt machine learning techniques for both understanding of correlation between various properties and also using this relation to impute missing property data of fuels.

For integration of advanced ML strategies, an effort is being made to convert much of the property data to a CSV format. In the future, analysis of the data using advanced techniques will also be available for download directly from the website using a similar format. The presentation of fuel data in a universal format will allow the information to be read by multiple software using an appropriate script to create tailored, reproducible output for specific user needs.

System Operations

As a consequence of improved access to specific fuels data, there are possibilities in terms of improving the quality control and tracking of fuels in use: through the processing and production of the fuel, through the supply chain and to its end use. The fuel data is currently transferred through the system using the quality assurance certificates, which then need to be manually integrated into fuel surveys or modelling tools. If the data were stored in a common electronic schema or format and processed into a quality document at the point of need, it would increase the flexibility of this data.

The availability of this data in an interoperable format will be of benefit for both airports and airlines. It will simplify the gathering of evidence to demonstrate the safe usage of SAF with detailed supply information. It is hoped that this will build trust between producers, suppliers and consumers. Finally, the common electronic schema for fuels data
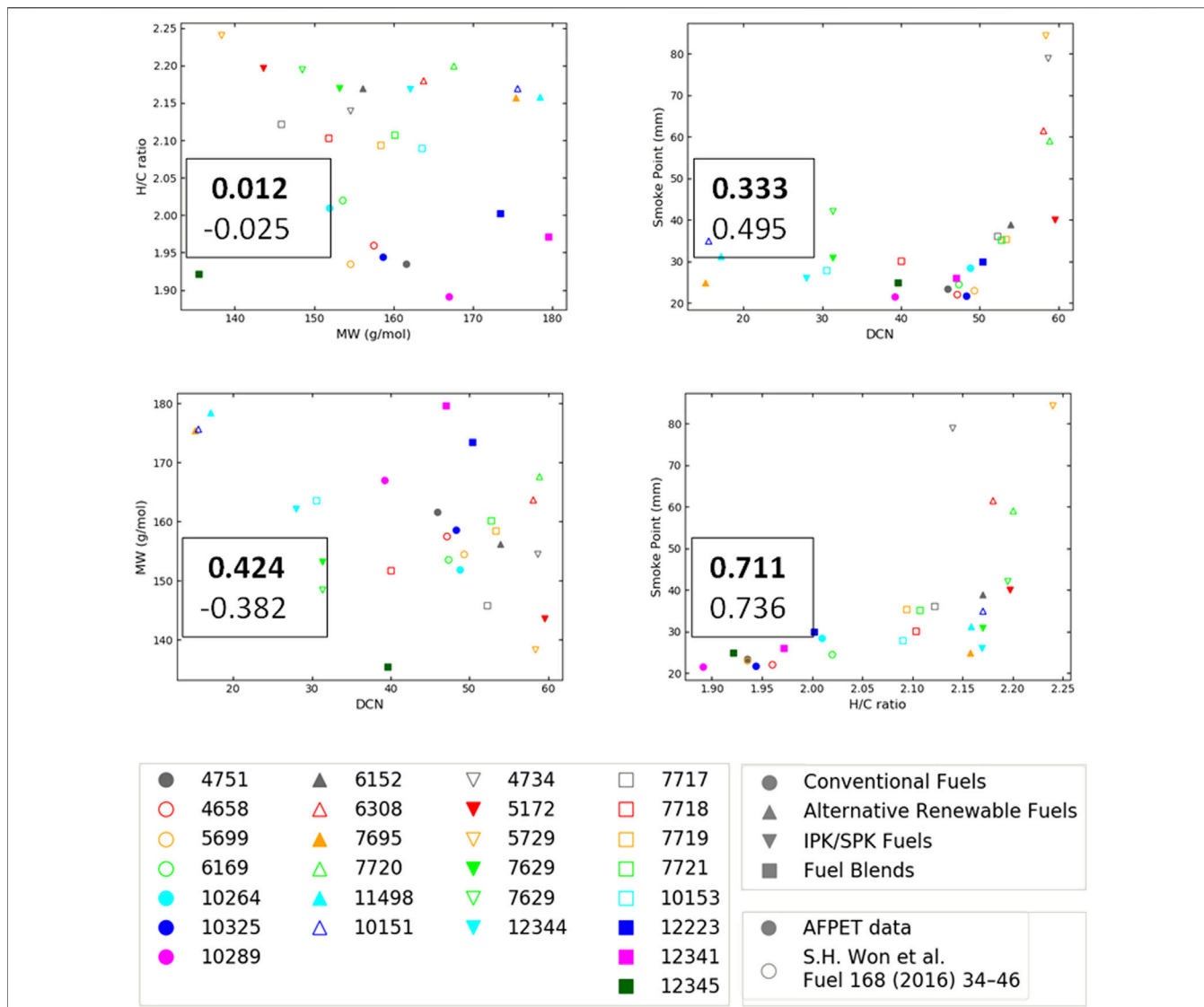
FIGURE 7 | Searching for correlations of fuel properties to emissions performance linking to specific fuel compositional families (indicating the R^2 correlation coefficient: bold type set indicates solid data points, regular type set indicates hollow data points), clockwise from top left, molecular weight vs hydrogen/carbon ratio, DCN vs smoke point, smoke point vs H/C ratio and DCN vs molecular weight.

throughout the supply chain will expand the availability of operational data and increase systems optimisation across the industry.

For just the fuel specification properties, the size of the required .json file is around 5 kB, which grows significantly if beyond specification fuel composition and performance properties are considered. Unfortunately this is larger than the data which can be stored within a QR code (3 kB), however key fuel properties could be stored within a QR code for a specific batch of fuel along with a hyperlink to a complete fuel .json file which could be considered a digital twin of the fuel, travelling through the fuel system. This type of fuel information would be of great usefulness for the community in the future. As it would

facilitate effective fuel blending (especially for high blending ratios) and informing the airline operator and pilots about the actual fuel in use.

Figure 8 shows the analysis of results from the METRON program in the U.S. (https://altjetfuels.illinois.edu/), which was a survey of fuels from the U.S. domestic airports as part of the ASCENT program. **Figure 8** shows the change in fuel properties with location, the trend in aromatics level across a single year as well as longer term trends at individual locations for specific fuel properties. For these results, it is clear that the overall average aromatic content is slightly less than 16%, which when blended at 50% with HEFA SAF would give a value to just satisfy the 8% minimum aromatic content requirement since aromatics content
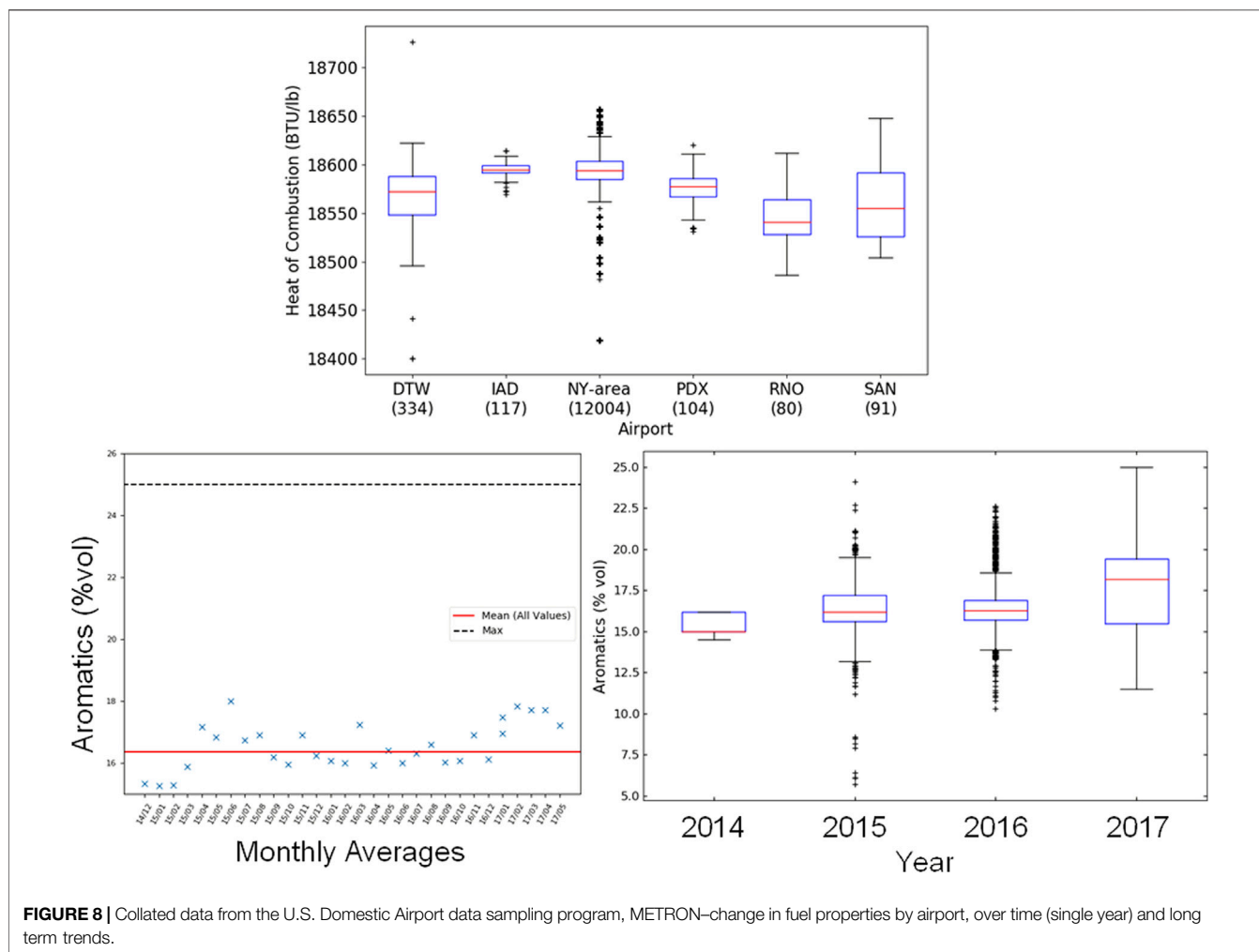
FIGURE 8 | Collated data from the U.S. Domestic Airport data sampling program, METRON–change in fuel properties by airport, over time (single year) and long term trends.

requirement since aromatic content shows linear blending properties.

This does also suggest that if sufficient volumes of SAF are available for higher than 50% blending, then higher blend ratios of SAF could be achievable for higher aromatic conventional fuels whilst still complying with the minimum aromatic content of 8%. This fuel would currently be out of the specification limits, whilst still complying with the fuel normal behaviour as suggested in **Figure 4**. With the currently available data and operational processes, it is most likely that fuel suppliers will be more conservative than this maximum blending ratio–for example, a 30:70 blend would ensure that the fuel always meets the specification, accounting for variability in the conventional fuel aromatic content assuming no information regarding the actual aromatic of the conventional fuel is available. If more knowledge of specific fuel composition were to be more readily available, say as part of the fuel supply then such a decision could be justified in the future. In order to reduce any risks associated with reaching the limits of the specification for this fuel property a more complete understanding of the aromatic components of the fuel is required. The composition of the aromatic portion of the fuel will also impact the fuel performance and is an area of current and future research.

DISCUSSION

The availability of specific fuel data is increasingly important in the current climate for aviation fuels. A significant amount of data is already available and is currently of a low level of utilisation due to the formatting of the data which lowers its impact. This is particularly the case for the data of fuel surveys which unfortunately means that it is possible for fuel related decisions are made without full knowledge of the current range of fuels in used in service and an over reliance on "average" fuel properties. This can be addressed by the methods presented in the current paper and this is presented with some usage cases.

In the future there is likely to be a further need for fuel specific data which can be provided by a Digital Twin of the fuel as it travels through the fuel system. For example, more detailed knowledge of the exact density, aromatic content and calorific volume of fuel uploaded onto aircraft can remove the need for calculating conservative estimates of fuel requirements for particular flight missions. This operational change could result in a reduction in fuel burn and consequently CO_2 emission during flight as well as

improved understanding of fuel impact on emissions and non-CO_2 impacts on a specific flight mission (Voigt et al., 2021). As the levels of SAF increase, it will become important for ascribing environmental impact from particular flights ultimately as part of the ICAO Carbon Offsetting and Reduction Scheme for International Aviation (CORSIA). Indirect changes in the CO_2 emissions of flight such as the example cited above due to the fuel properties will need to be taken into account to avoid the unnecessary consumption of aviation fuels.

Currently the CORSIA certification model does not have any fuel type specification details. Although this is perhaps not envisaged for the early stages of CORSIA, it is something that will be needed for the full monitoring and control of environmental impact as the levels of SAF increase. In addition, the non-CO_2 impact of the uptake of SAF will need to be properly accounted for.

This schema system is intended to be as flexible as possible for integration into an unstructured database for carrying out queries and searches of the database. The database can be extended to include fuel production information, where publically available, and specifically for environmental impact, the life cycle assessment performance of particular fuel production processes.

The reliability any conclusions drawn from the use of such databases is a function of the quality of information provided in the fuel database .json files. Much of the data is built on the fuel specification results of D1655 and DEFSTAN -091 91 for which experimental methods are standardised and whilst may not be as accurate as research methods for determining fuel properties, they are consistent and can be used for comparison across very large datasets. A positive case in point is the distillation standard D86, which is a simple distillation process which has known short comings, but is very repeatable across fuel laboratories and facilities. A less positive case would be the Smoke Point results, which have been shown to be highly stratified by the fidelity of the experimental set up in D1322, and influenced by the 25 mm cut off for the requirement of Naphthalene testing.

Nonetheless, the inclusion of performance data from fit for purpose or in service testing would be more susceptible to repeatability issues across different laboratories. Due to the unstructured nature of the .json format, it would be possible to ascribe meta data to any fuel property indicating a level of confidence in the recorded data. This would be essential to allow the data to be treated with an appropriate level of confidence.

There is an increasing desire to track individual fuels through the supply system particularly in terms of ensuring quality control as the fuel passes through the system. Having access to the specific fuel digital twin as well as the specification information of the fuels in the supply system at the same time would allow the assessment of any cross contamination or blending issues in a much more flexible and scalable manner than with previous systems.

In the current state of the databases, the overseers of information quality are the authors of this manuscript. As this database develops further, one approach would be for the data to be overseen by an independent body similar to those that already host fuel property data. Alternatively, the data could be maintained using a Block Chain approach, removing the need for a single central authority. Through the approaches the authors are making to share data across servers, the database is approaching the position of a leger of "blocks" of data across many servers which should ensure data integrity. However, careful consideration needs to be given to this step, were it to be taken.

CONCLUSION

The concept of a shared data schema allows fuels data to be far more flexible in the future. A feature that is important for a developing multi stakeholder fuel and SAF industry. This paper presents an interoperable and scalable method for the sharing of such data. This work initially was focused on supporting the early stages of fuel pre-screening, however it has quickly become apparent that the system proposed could make fuel by fuel data available to be utilised for a wide range of usage cases.

A graphical visualisation method to simplify the complex data in the fuel specifications into a format that can be easily interpreted by eye is presented, and can be used for communication of a range of otherwise complex comparisons between specific fuels and the average and specification limits of performance.

The supporting material provided with this paper provide an initial toolset for users to develop their own datasets and the authors would be encouraged by feedback or suggestions as to how the usage set could be expanded in the future.

It would be recommended that future publicly funded fuels research programmes adopt the schema for the recording of their fuels data as part of their data management plans. This will ensure the interoperability of data in future.

The next steps in developing this approach are sharing a common schema and establishing a platform for data sharing along with a protocol for the addition of similar databases in the future. There is significant demand for a public dataset particularly for SAF which the JETSCREEN and ASCENT projects will both be releasing in their own programmes. It is important to not to lose the benefit of conventional fuels data. The monitoring of the impact of the uptake of SAF is also an important process the database approach can support.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SB, TL, BR and AO contributed to conception and design of the study. SB, TL and BR organized the databases. SB, BR and AO performed the statistical analysis. SB wrote the first draft of the manuscript. SB, TL, BR, and AO wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

REFERENCES

altjetfuels (2015). Federal Aviation Administration National Alternative Jet Fuels Test Database. Available at: https://altjetfuels.illinois.edu/.

ASTM (2020). *D4054-20c, Standard Practice for Evaluation of New Aviation Turbine Fuels and Fuel Additives*. West Conshohocken, PA: ASTM International.

ASTM International (2020). *D1655-20a, Standard Specification for Aviation Turbine Fuels*. West Conshohocken, PA: ASTM International.

Colket, M., and Heyne, J. (2021). *Fuel Effects on Operability of Aircraft Gas Turbine Combustors*. Reston: AIAA. doi:10.2514/4.106040

Coordinating Research Council (1983). *Handbook of Aviation Fuels Properties*. Alpharetta, GA: CRC.

Coordinating Research Council (2014). *Coordinating Research Council Inc. ALPHARETTA, GA 30022, Report No. 663, Aviation Fuel Properties Handbook 4th Ed.*

Dryer, F. L., Jahangirian, S., Dooley, S., Won, S. H., Heyne, J., Iyer, V. R., et al. (2014). Emulating the Combustion Behavior of Real Jet Aviation Fuels by Surrogate Mixtures of Hydrocarbon Fluid Blends: Implications for Science and Engineering. *Energy Fuels* 28 (No. 5), 3474–3485. doi:10.1021/ef500284x

Energy Institute; QinetiQ (2014). *The Quality of Aviation Fuel Available in the United Kingdom Annual Survey 2014*. London: Energy Institute.

Grahe, J. E., Cuccolo, K., Leighton, D. C., and Cramblet Alvarez, L. D. (2020). Open Science Promotes Diverse, Just, and Sustainable Research and Educational Outcomes. *Psychol. Learn. Teach.* 19 (1), 5–20. doi:10.1177/1475725719869164

Hadaller, O. J., and Johnson, J. M. (2006). *Coordinating Research Council Inc. ALPHARETTA, GA 30022, Report No. 647, World Fuel Sampling Program.*

Hall, C., Rauch, B., Bauder, U., Le Clercq, P., and Aigner, M. (2021). Predictive Capability Assessment of Probabilistic Machine Learning Models for Density Prediction of Conventional and Synthetic Jet Fuels. *Energy Fuels* 35, 2520–2530. doi:10.1021/acs.energyfuels.0c03779

Heyne, J., Bell, D., Feldhausen, J., Yang, Z., and Boehm, R. (2022). Towards Fuel Composition and Properties from Two-Dimensional Gas Chromatography with Flame Ionization and Vacuum Ultraviolet Spectroscopy. *Fuel* 312, 122709. doi:10.1016/j.fuel.2021.122709

Heyne, J., Opacich, K., Peiffer, E., and Colket, M. (2019). "The Effect of Chemical and Physical Fuel Properties on the Approval and Evaluation of Alternative Jet Fuels," in 11th U.S.National Combustion Meeting, Pasadena, CA, March 24-27, 2009.

Heyne, J., Rauch, B., Le Clercq, P., and Colket, M. (2021). Sustainable Aviation Fuel Prescreening Tools and Procedures. *Fuel* 290, 120004. doi:10.1016/j.fuel.2020.120004

Ministry of Defence (2016). *Defence Standard 91-091, Turbine Fuel, Aviation Kerosine Type, Jet A-1, NATO Code: F-35; Joint Service*. AVTUR (9).

Moses, C. (2017). *AFRL-RQ-WP-TR-2017-0091, Delivery Order 0006: Airbreathing Propulsion Fuels and Energy Exploratory Research and Development (APFEERD), Subtask: Review of Bulk Physical Properties of Synthesized Hydrocarbon: Kerosenes and Blends.*

Oßwald, P., Zinsmeister, J., Kathrotia, T., Alves-Fortunato, M., Burger, V., van der Westhuizen, R., et al. (2021). Combustion Kinetics of Alternative Jet Fuels, Part-I: Experimental Flow Reactor Study. *Fuel* 302, 120735. doi:10.1016/j.fuel.2021.120735

Rauch, B. (2020). JETSCREEN: JET Fuel SCREENing and Optimization 2020. Available at: https://cordis.europa.eu/project/id/723525.

The Defense Energy Support Center (2009). *Petroleum Quality Information System Annual Report 2008*. Fort Belvoir: The Defense Energy Support Center.

Voigt, C., Kleine, J., Sauer, D., Moore, R. H., Bräuer, T., Le Clercq, P., et al. (2021). Cleaner Burning Aviation Fuels Can Reduce Contrail Cloudiness. *Commun. Earth Environ.* 2, 114. doi:10.1038/s43247-021-00174-y