Characterization of Agricultural Aircraft Performance Using Flight Log Data

MISSISSIPPI STATE
UNIVERSITY™

RASPET FLIGHT RESEARCH
LABORATORY

July 2020

Table of Contents

DISCLAIMER

This analysis, and the resulting example flight profiles, are derived from source data that is wholly owned by MSU, and was collected outside of any government-sponsored effort. The source data is subject to multiple non-disclosure agreements, and thus is not directly disclosed in this report. MSU has reviewed the enclosed summary to ensure that it 1) accurately reflects our source data regarding typical flight behavior of sampled aerial application aircraft; and 2) is releasable to the government for the express purpose of supporting project *A18_A11L.UAS.22 Small UAS Detect and Avoid Requirements Necessary for Limited Beyond Visual Line of Sight (BVLOS) Operations: Separation Requirements and Training*. No other use or release of this information is authorized without the expressed written consent of RFRL.

Revision number	Edition date	Reason for change	Pages/Sections Affected
20200630-RFRL-AG DATA-1006 REVISION 1	06/30/20	First draft	none
20200720-RFRL-AG DATA-1006 REVISION 2	07/20/20	General edits and formatting	All
20200721-RFRL-AG DATA-1006 REVISION 3	07/21/20	Edits to Disclaimer	3
20200825-RFRL-AG DATA-1006 REVISION 4	08/25/20	External Edits	All

1 Introduction

The integration of Unmanned Aerial Systems (UAS) into the National Airspace (NAS) poses a risk to both the current population of manned aircraft and the future addition of UAS both large and small. In order to mitigate the safety risks, the Federal Aviation Administration (FAA) has called for diverse research on a multitude of subjects including the establishment of commercially and federally accepted methods of characterizing and measuring risks, characterization of the current NAS, planning for the future integrated NAS and many other related topics. One special topic is the development of an operational model for aerial application of dispersants over agricultural fields and the expected behavior of the aircraft in rural and farmland areas. The aircraft that populate this special case behave unlike the General Aviation (GA) and private pilots that make up suburban and urban airspace. These aircraft fly significantly closer to the earth's surface than GA and operate with precision systems that assist in lining up approaches to a field. Their rate of climb and descent are a function of their position relative to a field they will spray with liquid or dry dispersants. Generally, an agricultural aircraft will make a handful of landings a day to refuel and refill their payload. The unique behavior of this specific genre of aircraft must be quantified and made into models that contribute to the movement toward safe introduction of UAS into shared airspace with manned traffic. This report covers the acquisition of agricultural aircraft flight logs, the characterization of that data set, the methodology behind quantifying aircraft behavior, the resultant statistics associated with such and the proposal for further and more in-depth research.

1.1 Call for Participation

In 2017, Raspet Flight Research Laboratory (RFRL) asked the National Agricultural Aviation Association (NAAA) for volunteer groups of agricultural pilots, or operators, to submit their flight log data and be a part of the growing movement toward UAS integration. After a few years, a total of about 30 operators submitted 30,000 flight logs and the privacy of the operators and their pilots were preserved by removing any names attached to the data. The NAAA noted that a condition of participation was restriction of distribution and use of the data outside of the bounded agreement.

1.2 Performance Trends

The five most important aircraft behavioral trends for designing systems to safely operate around agricultural aircraft are discussed in this section. The angles that these aircraft tend to fly relative to the Earth's surface will help guide the testing of Detect-and-Avoid (DAA) equipment as the resolution of an intruding aircraft may be lowered by sharp angles between the DAA system and the target aircraft. Spray speed and the altitude at which these aircraft fly when applying pesticides to crops will give insight into what altitude a DAA system's lower boundary must capture. Last, the cruise performance in terms of speed and altitude highlights the difference between lower airspeed GA and higher performance aircraft including agricultural aircraft. The following subsections provide detailed expectations about these behaviors.

1.2.1 Angles Relative to Earth's Surface

As an agricultural aircraft descends, onboard flight instruments assist the pilot in maintaining the aircraft's altitude above crops, trees, and possible obstructions such as power lines. The pilot will also attempt to level off the aircraft prior to entering the volume above a field. When the pilot finishes a spraying run, the aircraft will ascend more rapidly to a comfortable altitude as the airspace around the field will usually be less obstructed beyond the tree line. Therefore, the expectation is the average angle of descent should be noticeably less than the angle of ascent.

1.2.2 Spray Speed and Altitude

After descending from cruise or in between spraying runs, the agricultural aircraft should see large acceleration as the potential energy trades off for higher airspeed. To effectively spread the payload, the aircraft will need to be low to the earth while passing fields of crops. Considering a survey of different agricultural aircraft models as well as review of archived flight footage, the mean agricultural aircraft spray speed should be around 120 knots and the altitude less than 50 feet.

1.2.3 Cruise Speed and Altitude

One important note about this metric for aircraft behavior lies in the fact that agricultural aircraft can be expensive to operate. Private businesses attempt to optimize flight plans for minimizing fuel, maximizing spray areas covered in one flight and quickly finishing a job. Noticeably in the data set, these aircraft tend to reach a peak altitude then somewhat immediately begin a descent toward the first field. Simply, long periods of cruising at a consistent speed and altitude uncommonly appear in the data set. Therefore, the method for analyzing this inconsistency as well as the future design on safety cases for UAS operation must take this into account. The expected trend for cruise speed should be slightly lower than the average cruise speed in the following section's survey since these aircraft may never operate at the constant altitudes used to define their aircraft's performance (above 4,000 ft MSL).

1.3 Survey of Agricultural Aircraft

To compare the model generated with the provided data, a brief survey of existing agricultural aircraft was performed. Knowledge of the specifications of the different agricultural aircraft models would help ensure the model parameters found are reasonable and within expected ranges. Not only do current aircraft specifications help with validating model parameters, it also could help identify performance trends within companies and potentially explore aircraft classification. Table 1 below shows an overview of agricultural aircraft from American manufacturers with various performance specifications.

Table 1. Survey of Agricultural Aircraft Specifications.

Aircraft Name	Spray Speed (mph)	Rate of Climb (ft/min)	Cruise Speed (mph)	Stall Speed – unloaded (mph)	Stall Speed – loaded (mph)
Weatherly 620-B	129	700	115	54	65 (5000 lbs)
Weatherly 620-BTG	140	1400	-	-	71 (6000 lbs)
Air Tractor AT-402B	120-140	800	162	53	66 (flaps up) /77 (flaps down) (7,000 lbs)
Air Tractor AT-502B	120-145	870	154	53	68 (flaps up) /82 (flaps down) (8,000 lbs)
Air Tractor 502XP	120-140	816	167	53	68 (flaps up) /82 (flaps down) (8,000 lbs)
Air Tractor AT-504	120-145	860	151	53	65 (flaps up) /75 (flaps down) (8,000 lbs)
Air Tractor AT-602	145	650	182	60	82 (flaps up) /99 (flaps down) (12,500 lbs)

1.3.1 Spray Speed

Crop dusting involves spraying crop protection products on crops from agricultural aircraft. As crop protection products tend to be liquid in nature, a range of values applicable to all aircraft is expected to ensure these products are uniformly applied on all crops. Figure 1 below shows the distribution of spray speeds for agricultural aircraft.

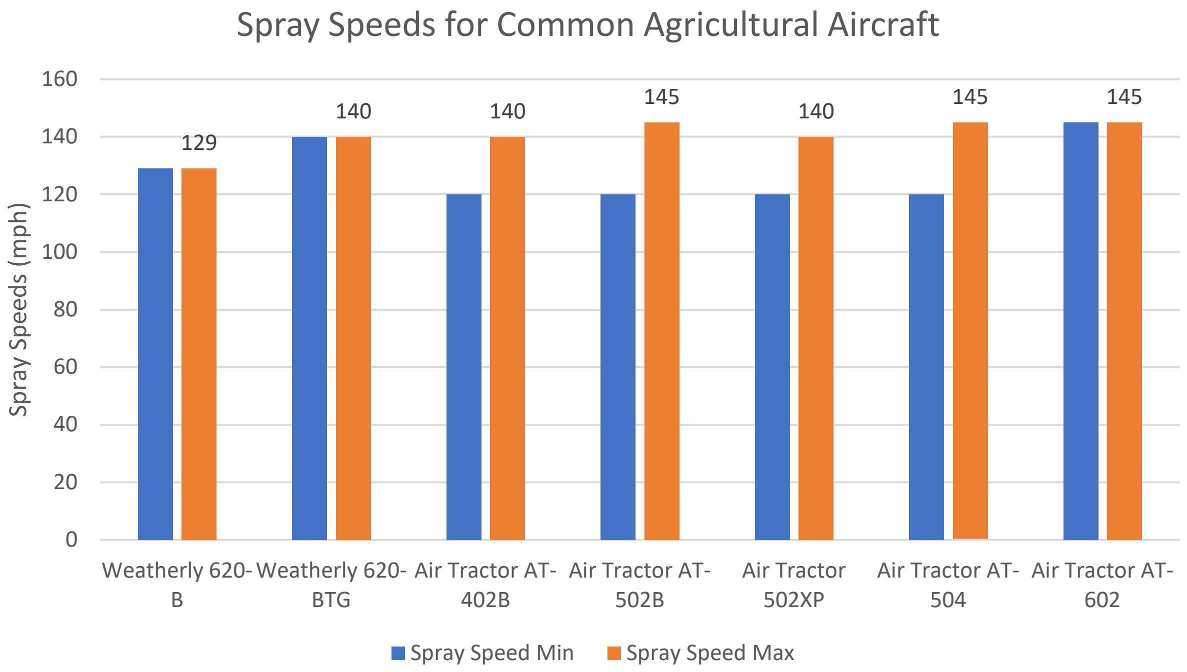


Figure 1. Spray speeds (mph) for common agricultural aircraft.

1.3.2 Cruise Speed

The cruise speed of agricultural aircraft varies based on the size and dimensions of the aircraft. This value represents the airspeed of the aircraft, which can be incomparable to ground speed on highly windy days. Figure 2 shows the cruise speed for the aircraft types listed above.

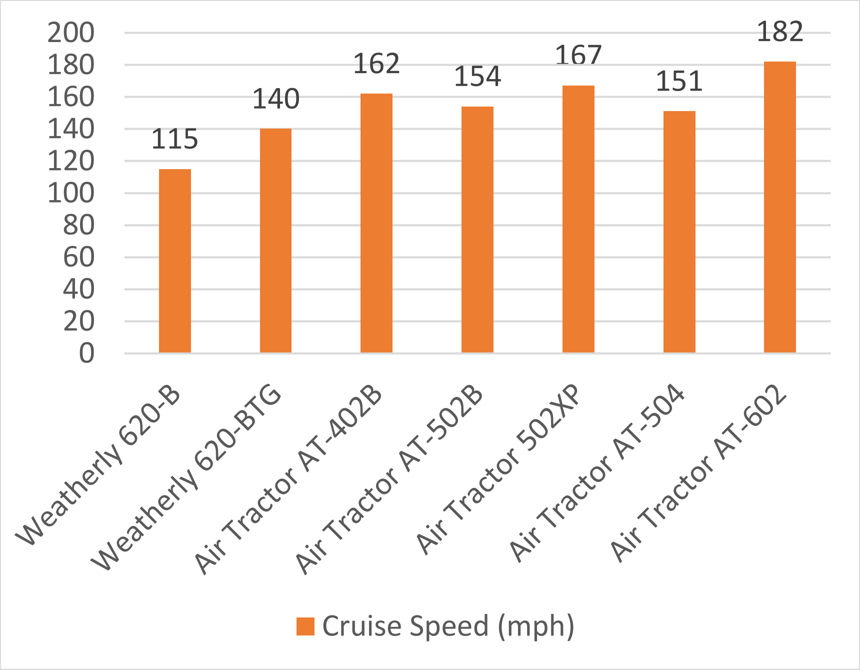


Figure 2. Cruise speeds (mph) for common agricultural aircraft.

2 Data Characterization

Every flight log has latitude and longitude coordinates corresponding to its location during flight time. Each state has politically defined borders also represented as a series of latitude and longitude coordinates and is used to determine which state each flight log was taken from. All flight logs will have coordinates that exist only inside one state at a time; by characterizing which state each log belongs to, we can obtain a physical count of how many logs reside inside each state. Figure 3, shown below, shows the regions the data was collected from.

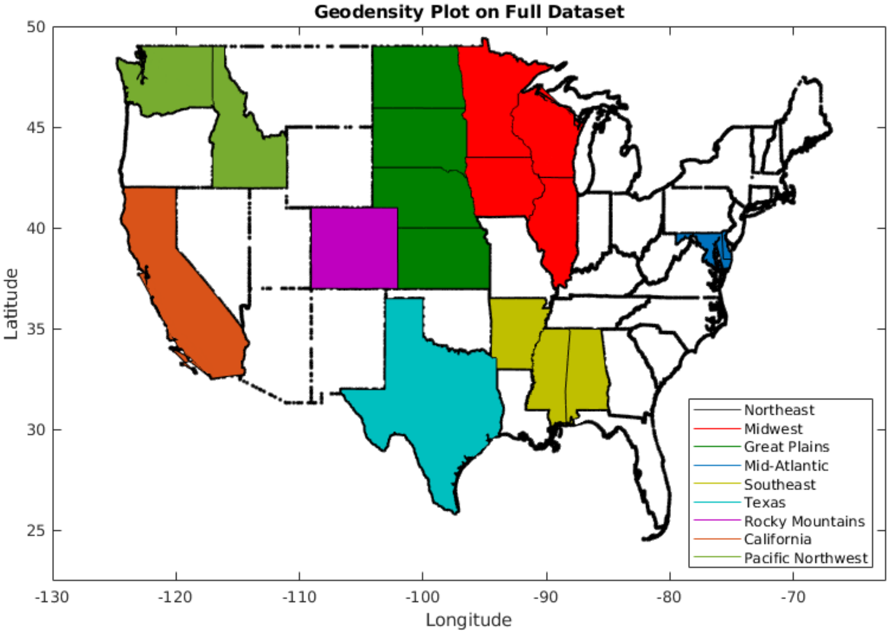


Figure 3. Defined regional clusters.

Constructed out of a total 28,626 files, Figure 4 below visualizes the flight log population geographically for the dataset. A color closer to red indicates that the state holds a higher percentage of logs in the dataset. In this dataset, there are four states that comprise approximately 70% of this dataset: Minnesota, Illinois, Idaho, and Texas. Additionally, there are seven more states that represent at least 2% of the dataset (>562 files): North Dakota, South Dakota, Nebraska, Iowa, Arkansas, Alabama, Mississippi, Maryland, and Delaware. This indicates the given dataset is primarily skewed towards the Great Plains/Midwest regions, the Pacific Northwest, and Texas. There are 9 files outside the scope of the Continental United States, with 1 log located in Alaska and 8 logs located outside the United States. These files were not included in the final statistics.

By classifying these states into different "regions," clustering can be performed to represent which regions of the country have flight logs. From the histogram in Figure 5, a significant amount of flight logs exist in the Midwest United States (Minnesota, Iowa, Illinois, Wisconsin, Michigan) as expected, with more flights in the Pacific Northwest (Washington and Idaho), the Great Plains (North/South Dakota, Nebraska, Kansas, Oklahoma), and Texas.

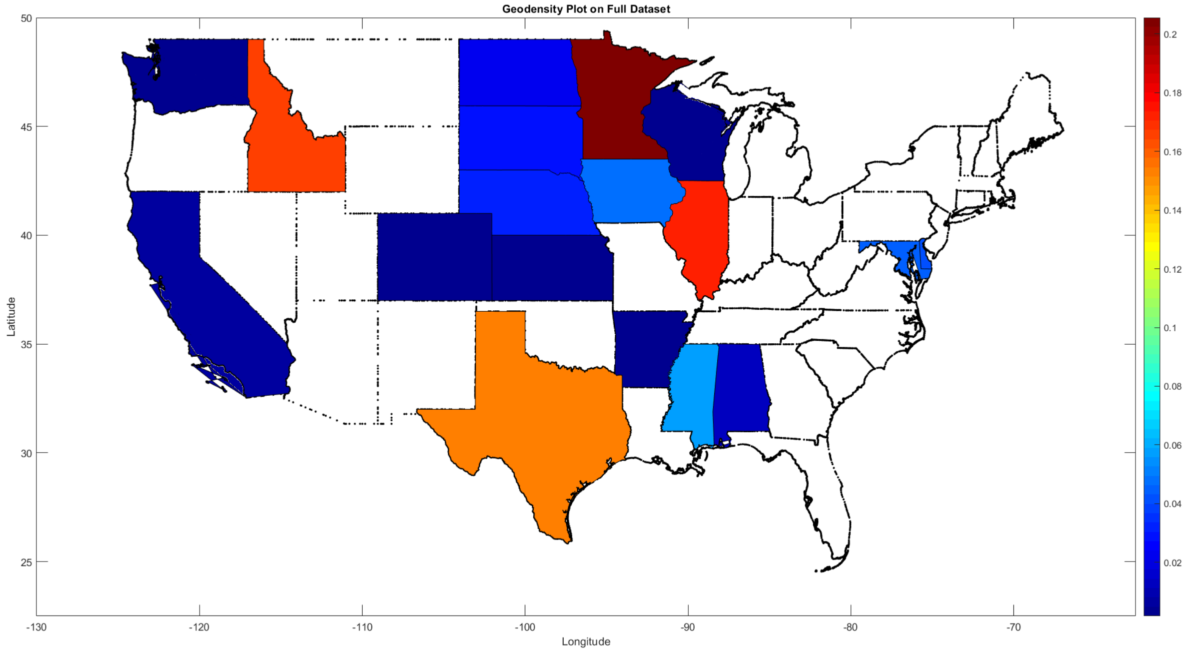


Figure 4. Flight density plot of the continental U.S. Color represents number of files as a fraction of the whole data set.

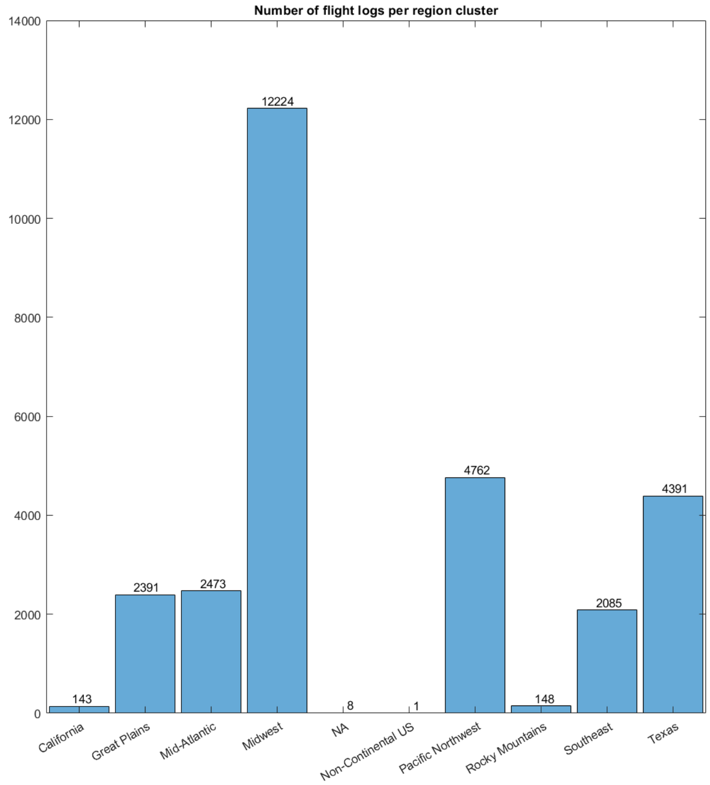


Figure 5. Number of flight logs per region cluster.

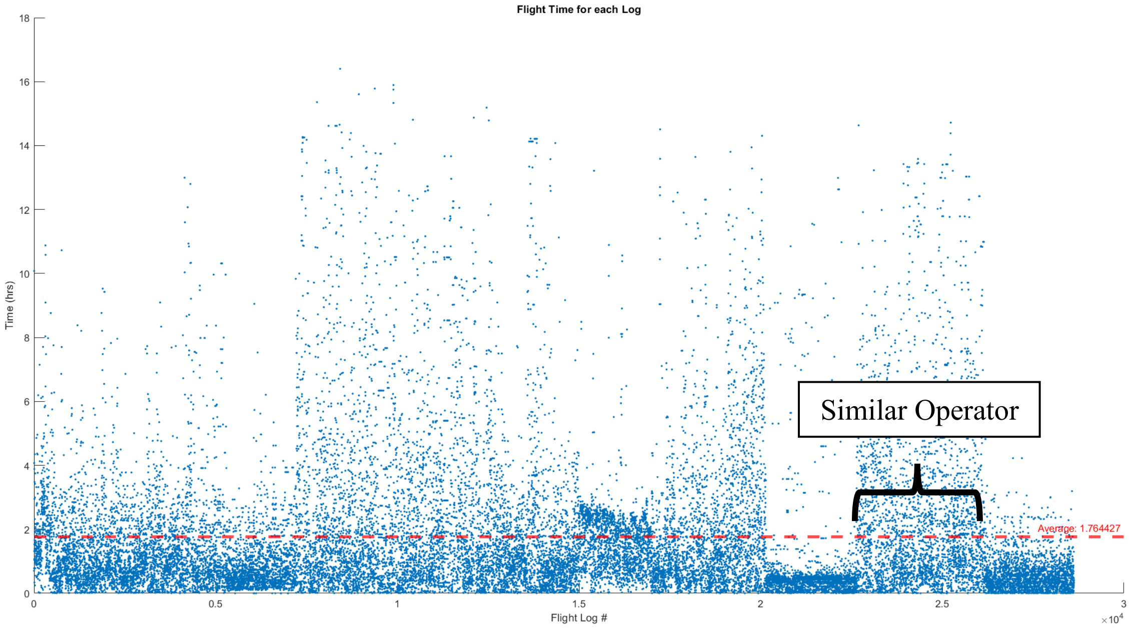
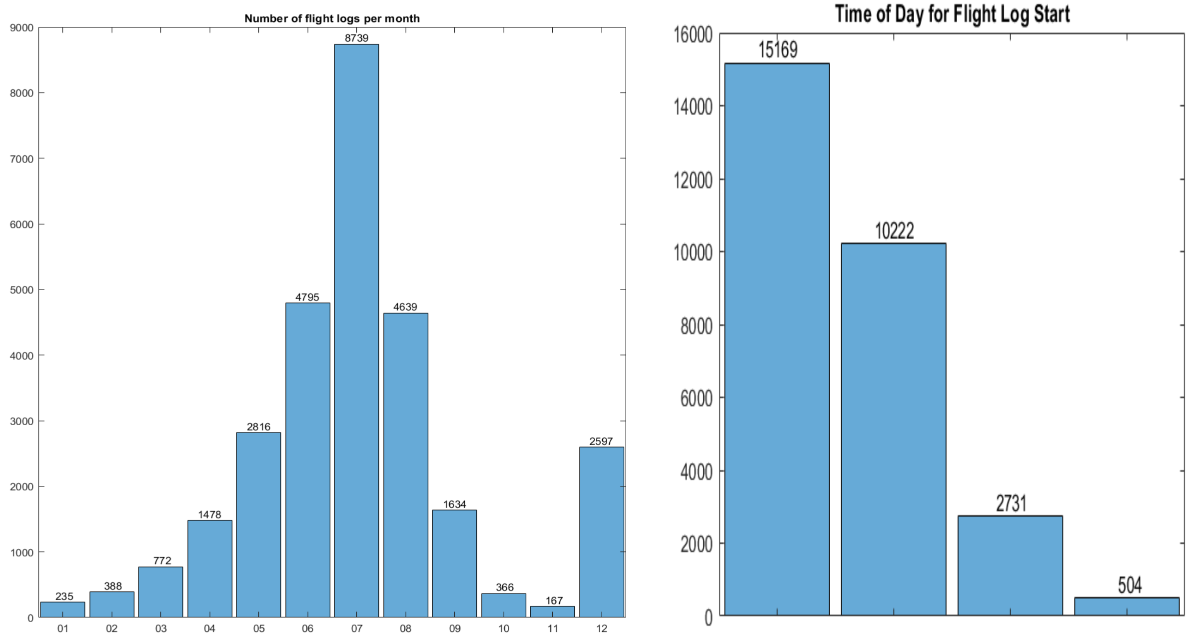


Figure 6. Average flight log length with average of 1.76 hours.

Figure 7 visualizes the number of flight logs per month given a January to December or 1 to 12 calendar year and Figure 8 shows the most populated time of day for this data set.



Figures 7 and 8. Number of flight logs per month (left) and time of day (right).

2.1 Data Captured by Log

Each flight log consisted of several columns of possible data including ownship GPS position, spray valve state in terms of on or off, values for payload spray rate and the log metadata such as time and date. Less than 0.5% of the data contained entries for the laser altimeter's sensed altitude above ground level (AGL). The telemetry, such as angular rates and barometric altitude, of the aircraft was not included in the flight log data. This led to the development of several correcting and formatting procedures as outlined by the following sections.

2.2 Data Correction

After the flight logs were converted from a '.log' file extension to a usable format such as '.csv' or '.txt', some corrections and outlier removing was needed. A handful of the flight logs were outside of the continental U.S. for various reasons. These logs were removed for irrelevance. Other reasons for removal, correction, or consideration are described in the following sections.

2.2.1 Split Flights

Inherent to the field of agricultural aircraft is the need for constant refueling and therefore the splitting of operations and flights. A file with split flights may have long periods of time with 0 velocity or very little change to altitude and position. A split flight file may even be comprised of flights over a few days appended to one another. To tackle the issue of split flight data, a batch script was made to reorganize and rename these files into multiple new files. The resultant files contained corrected data and were then used in the final data set.

2.2.2 Formatting

Although the data was mostly uniform, two similar GIS software suites were used for translating the '.log' extension into a '.csv'. MapStar and NavViewW were used for this process. By manually selecting what properties of the original log the new '.csv' file should contain, a universal file type was determined and applied. The following eight major columns were converted first: date, time, latitude, longitude, altitude, speed, spray state, and AGL. It is worth noting that less than 100 files contained a value for AGL. The remainder of the software's possible outputs started in the ninth column continuing onward.

2.2.3 Geoid and Ellipsoidal Conversions

The process of converting Global Positioning System (GPS) data to AGL requires a process of converting from one model of the earth to an in-situ measurement that inherently contains some error. It was assumed the error for this in situ altitude measurement was small given the flat and plain-like nature of agricultural and arable land. See Appendix A for further details on the translation of GPS data to AGL.

2.2.4 Meaningful Data

The variety of log length, starting position within an operation, and other miscellaneous factors required the development of an acceptable file prerequisite. Flight logs consisting of 40 or less data points were considered not meaningful and were not included in the final data set. Duplicate files also did not add to the set. See Appendix B for further visualization of useful and not applicable files.

2.3 Usable Data

In order to easily manipulate the format of the data, MATLAB was used to create a customizable class with a list of properties that represent the columns of data. Each flight log was assigned as an object and appended to an array of all the file objects.

2.3.1 MATLAB Object Array

Each data file contained extensive information outside of what was required for the scope of this analysis. In order to easily work with the data, a class was defined that only contained the properties that were deemed immediately relevant. These were the original file path and name, the file's date, the start timestamp of the file, the time associated with each data entry, altitude, latitude, longitude, speed, spray state, and AGL in column order. The class also had properties for clustering; specifically, a state property, a geocluster property, a month property, and three generic identity properties for future applications. Once the class was defined, the data files were read into an object array and the whole dataset was saved as a MATLAB data file. Once the object array containing the whole data set was created, sub arrays were divided such that each contained an indiscriminate group of 2,500 files for debugging purposes. These steps were done to guarantee a future proofing of the data set as well as facilitate the appending of new flight logs.

3 Methodology

The development of a methodology for each key performance indicator included multiple steps. For those requiring an altitude AGL measurement, many calculations for evaluating performance as well as the accuracy of the assessment were required. The methodologies are a result of many iterations of increasing the accuracy of a performance trend. The following section includes the higher-level description of the assumptions taken to evaluate the trends, definitions of some key terms, and finally the methodologies behind the angles, spray characteristics and cruise characteristics behind the results section. Appendix C covers the following methodology sections in greater detail.

3.1 Angles Relative to Earth's Surface

The following subsections briefly cover the methodology behind analyzing the angles that the aircraft fly at while approaching and leaving a field. The rate of turn between runs is also included. Methods 1.1 and 1.2 describe the ascent and descent angles while method 1.3 deals with the rate of turn between runs.

3.1.1 Method 1.1

The ascent away from and descent to fields can be calculated using the spray valve state information. By knowing when the aircraft begins to operate over a field, this algorithm takes a window prior to the opening of the valve and another window shortly after closing. The windows are calculated using a static value for moving forward the index before the spray run and away from the spray run. This static method would set the baseline for further iterations of more precise methods. Table 2 visualizes an example column of data in row format. The calculations for angles depended on the change in altitude and ground position over time using trigonometric relationships.

3.1.2 Method 1.2

The second method for calculating angles uses a more reliable dynamic window. This window is determined by similar sequential changes in altitude that meet a certain cutoff criterion. The cutoff in Table 2 is at +/- 5 feet. This forces the window to only contain the sharpest of angles before and after a spray run and not any leveling off prior to or after a run. The results from this method further improved on method 1.1.

Table 2. Example file with descent in green and ascent in blue. The cutoffs for the algorithm defined in orange and yellow.

Delta Spray	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Delta Altitude	-1	1	-2	-5	-6	-4	-5	-6	-2	-1	-1	0	1	2	3	2	1	3	5	4	5	6	9	4	5	6	4	3	-1	1

3.1.3 Method 1.3

The idea for taking the turn rate of each flight log consists of evaluating the angles that each pilot takes for every turn between spraying runs. To achieve this, the periods of interest are defined as the times between the end of the previous spray and the beginning of the next spray. As this value is an instantaneous value, the average is determined by taking the mean across all turns, and an average per flight log is taken by subsetting each flight log as its own value. It is important to note that this method calculates a change in trajectory on a 2-dimensional scale as described in Appendix C.

3.2 Spray Speed and Altitude

The processes for determining the speed and altitude at which the aircraft flies while dispersing a payload over a target field are described in the following subsections.

3.2.1 Method 2.1

To determine the average speed, all entries in a data column for speeds that were between changes in state of a valve were considered. From the opening to closing of the spray valve, each value for speed was added to the list of averaged values to approximate a mean spray speed.

3.2.2 Method 3.1

Appendix A details the conversion from GPS altitude to an approximation for mean sea level (MSL). The difference between methods 3.1 and 3.3 is the model used for that process. In this method, the GMTED2010 and EGM2008 models are used. This conversion was only done for the altitudes while the spray valve state was open. The resulting altitudes were then averaged to approximate the mean spray altitude.

3.2.3 Method 3.2

The following method uses a subset of the larger data set. The logs added to this subset started with a low ground speed and a very small change in altitude. These attributes are similar to those of a takeoff or landing. The file's altitudes were biased using the aircraft's altitude while it is assumed to be still close to the ground on takeoff. This method produced similar results to method 3.3 as can be seen in the results section.

3.2.4 Method 3.3

Method 3.3 is identical to method 3.1 but uses the more recent Geoid12B geopotential model instead of the EGM2008 model. The values for altitude were only accepted if above -500 and below 500 feet.

3.3 Cruise Speed and Altitude

The following methods similarly addressed the challenges associated with converting GPS altitude to an approximation. Methods 4.1 and 4.2 cover the mean cruise speed while 5.1 and 5.2 approximate mean cruise altitude.

3.3.1 Methods 4.1 and 5.1

Methods 4.1 and 5.1 for cruise involve an intense process for clustering jobs with areas represented by polygons then taking the mean statistics while the aircraft travels to, from, or between n the clustered polygons. See Appendix C "Cruise Speed and Altitude" for more details.

3.3.2 Methods 4.2 and 5.2

These methods similarly cluster jobs but additionally provide a dynamic cutoff as described in Appendix C

4 Results

The following section provides results to the previously stated methods in the form of a distribution. Tables include the mean and standard deviation of each method's output.

4.1 Angles Relative to Earth's Surface

Tables 3 and 4 show the resultant angles and distributions for methods 1.1, 1.2 and 1.3. The results for method 1.1 are poor due to the use of a statically defined window. This static number could include invalid data that captures spray runs before or after an ascent or descent. Method 1.2 with various cutoffs provides more valid results.

M1.1 5-10 index:	Descent: -1.2526	
	Ascent: 0.6065	
M1.2 1 ft change cutoff:	Descent: -6.2113	σ = 1.2486
	Ascent: 5.5085	σ = 1.2877
M1.2 2.5 ft change cutoff:	Descent: -7.1934	σ = 1.3892
	Ascent: 7.0777	σ = 0.9986
M1.2 5 ft change cutoff:	Descent: -8.2003	σ = 1.5964
	Ascent: 8.4350	σ = 0.9992

Table 3. Ascent and descent angles (degrees) out of a spray run in degrees.

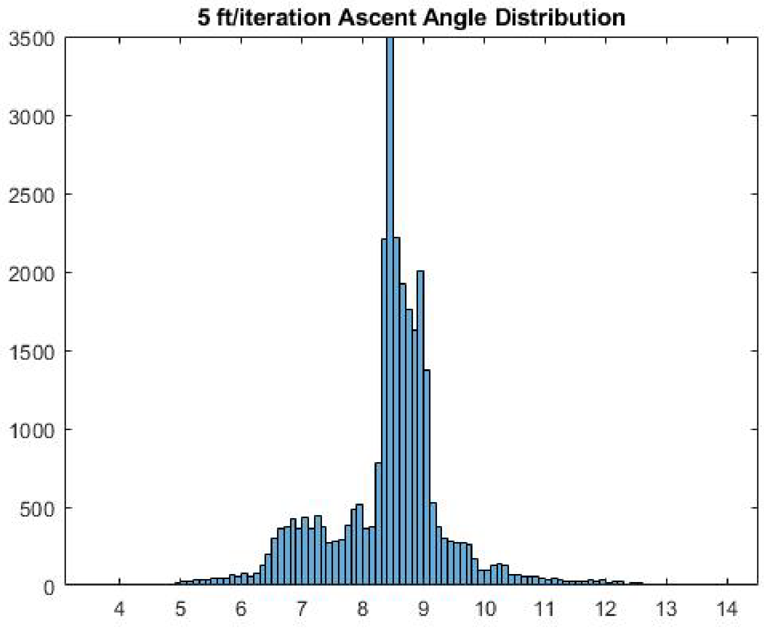


Figure 9. Angle of ascent (degrees) out of a spraying run distribution. Number of files on vertical axis.

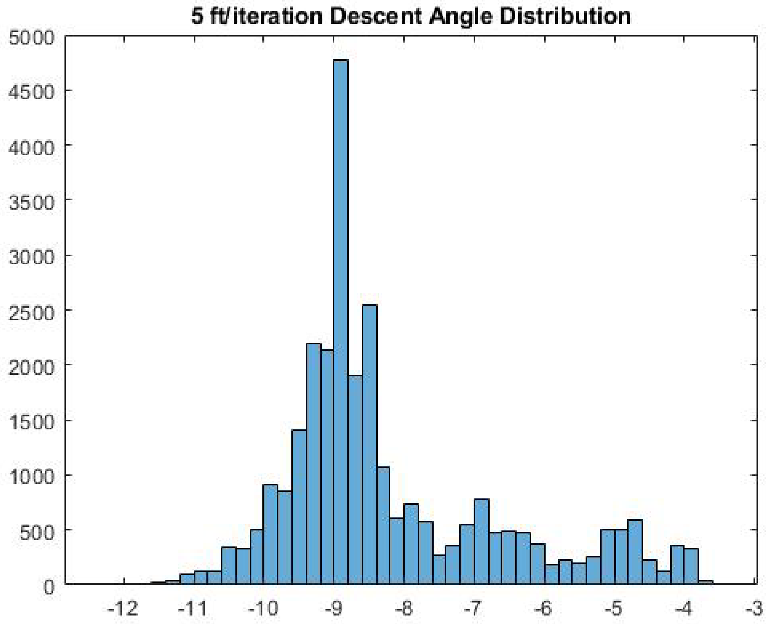


Figure 10. Angle of descent (degrees) out of a spraying run distribution. Number of files on vertical axis.

The turn rate was evaluated on two levels. Each individual flight log's turn average was then averaged all together. This could produce bias as pilots with longer flights and many more turns into and out of spray runs could be recorded, but only one average per file was created. The individual turn, approximately 1,973,107 turns in the data set, were averaged with equal weight. Both produced similar standard deviations of around 2.5 degrees.

Parameter	**Turn Rate Average per Flight Log**	**Turn Rate Average per Turn**
# data points	28256	1973107
mean	6.3357	6.831
STD	2.4502	2.6812

Table 4. Rate of turn in degrees.

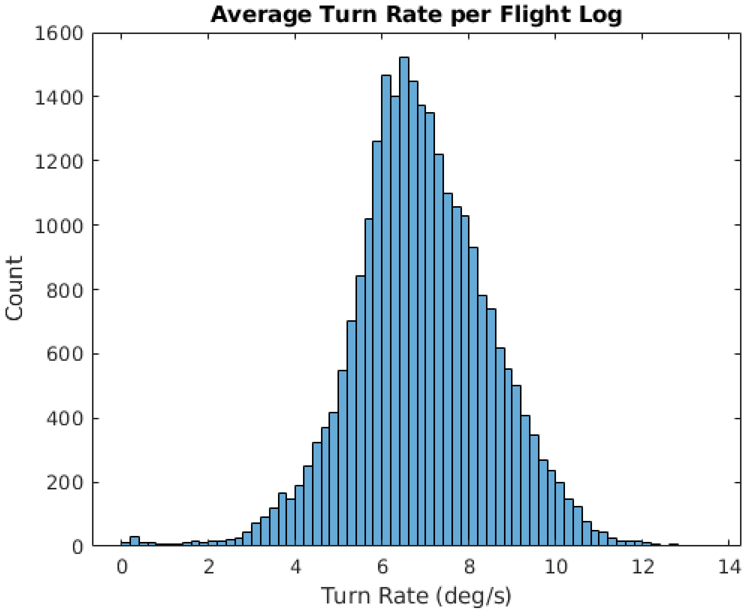


Figure 11. Average turn rate (degrees per second) between spray runs distribution. Number of files on vertical axis.

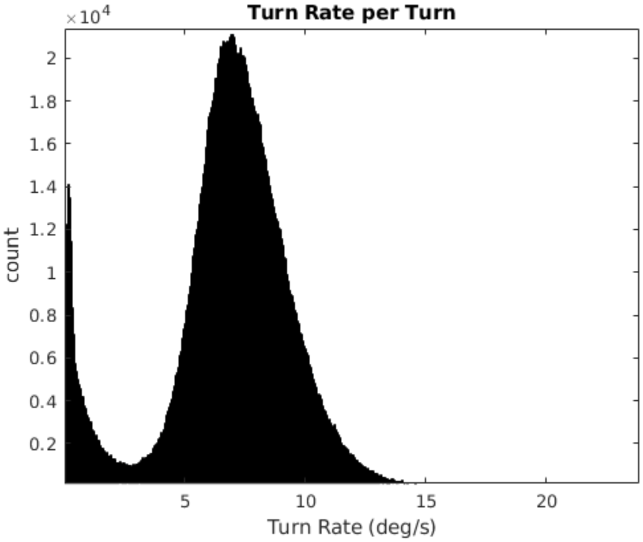


Figure 12. Average turn rate (degrees per second) between spray runs distribution. Number of data points on vertical axis.

4.2 Spray Speed and Altitude

The results in Table 5 should be very accurate as the method for obtaining them was simple and contained no approximations. The deviation seen could be the result of capturing a lower average for a spray run that requires the pilot to slow down due to obstacles or change in elevation.

M2.1:	Speed in mph: 139.48 Speed in knots: 121.21	σ = 17.4577 σ = 15.1705

Table 5. Speeds during a spraying run.

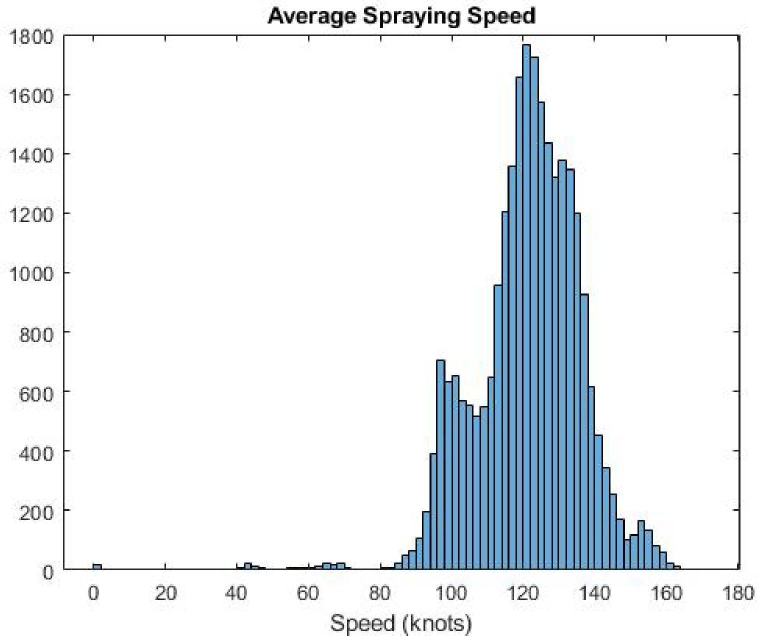


Figure 13. Average spraying speed (knots) distribution. Number of files on vertical axis.

M3.1 on Takeoff Dataset:	Altitude AGL in ft: 43.4434	
M3.2 on Takeoff Dataset:	Altitude AGL in ft: 34.3799	
M3.1 on Complete Dataset:	Altitude AGL in ft: 43.0787	σ = 145.7520
M3.3 on Complete Dataset:	Altitude AGL in ft: 37.9874	σ = 40.7703

Table 6. Altitudes during a spraying run.

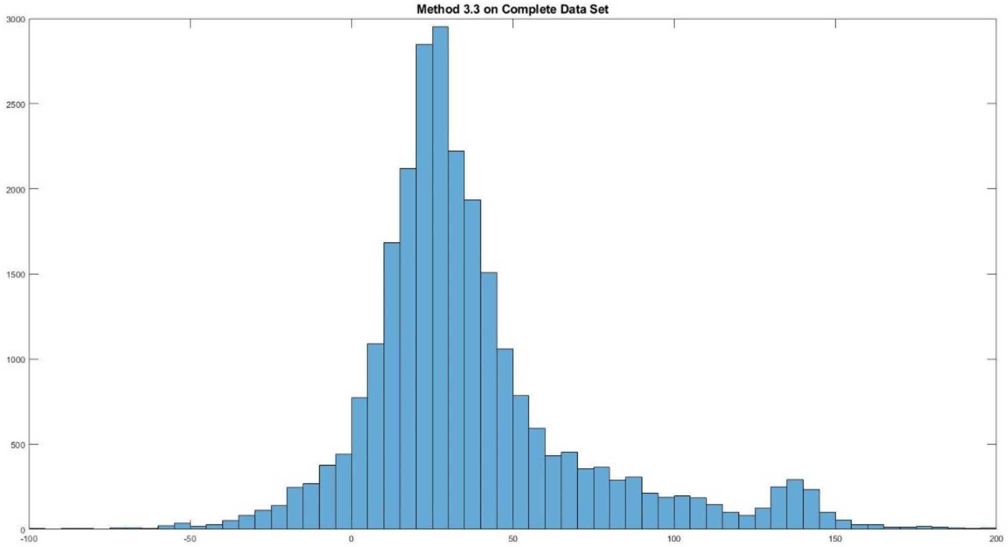


Figure 14. Spray altitude (feet) distribution. Number of files on vertical axis.

Even though each method for capturing spray altitude seemed to be an improvement in geoidal or ellipsoidal modeling, the results for mean spray altitude are similar. The extreme deviation seen in method 3.1 may be the effect of error in the elevation models as described in Appendix C.

4.3 Cruise Speed and Altitude

The cruise speed and altitude methods produced results within the performance specifications provided by the survey of common agricultural aircraft. The large deviation in cruise altitude between fields and airport is the product of the uniqueness of each file. Not every file had similar distance from airport to field or even time spent between fields. It is possible that the longer the time between fields the higher the altitude or longer time at cruise. More work can be done to organize the data set by distance between fields, from airport to field, and last field sprayed to airport.

M4.2	Speed in mph: 132.75 Speed in knots: 115.33	σ = 18.1198 σ = 15.7427

Table 7. Speed during cruising flight.

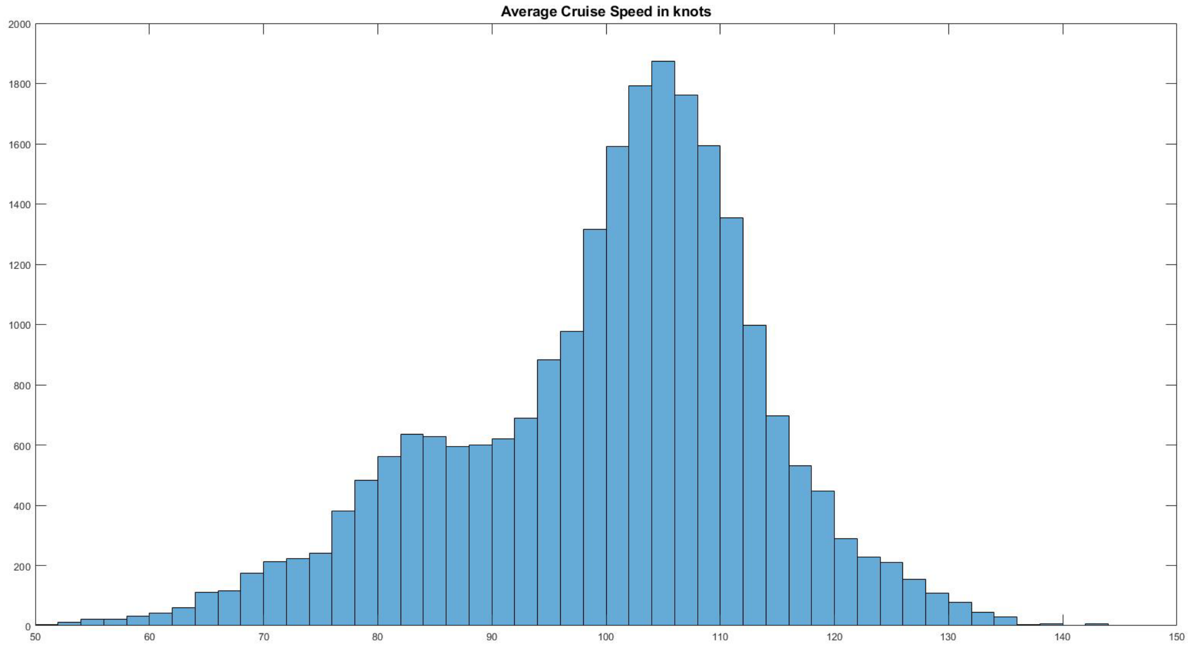


Figure 15. Average cruise speed (knots) distribution. Number of files on vertical axis.

M5.2	Altitude AGL in ft: 468.62	σ = 235.6166

Table 8. Altitude during cruising flight.

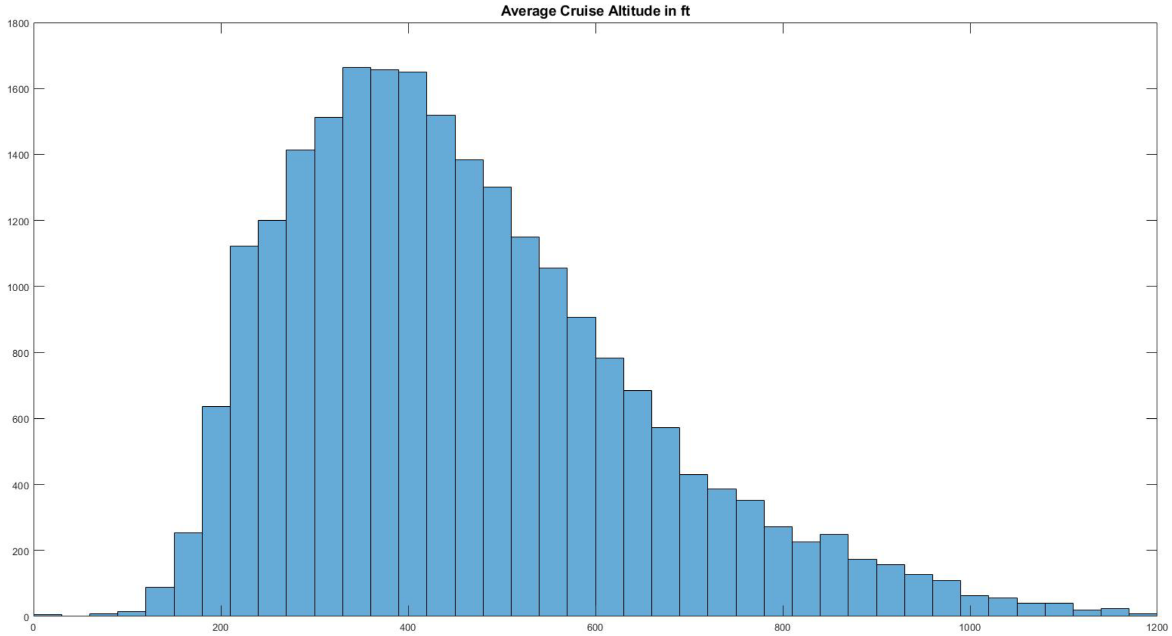


Figure 16. Cruise altitude (feet) distribution. Number of files on vertical axis.

5 Limitations

Although the data set consists of a large variety of operators, pilots, and locations, there exists room for improvement. By appending more flight logs to the data, the data set could be evaluated for more specific geographical and seasonal performance models as well as pilot behavioral models. The following subsections detail the possibilities for slight improvement of the data set.

5.1 Geographical Diversity of Participants

A large portion of the data set is set in the Midwest United States over the span of 2013 to 2017. By adding more flight operational logs to the data set outside of that region, the model could improve with respect to geographical diversity. This would help with distinguishing the probability of a crop type that the aircraft are spraying which could in turn affect the speeds and altitudes at which the aircraft spray.

5.2 Diversity of Log Date

The models in this report, as previously stated, are largely based around flight logs dated 2013 to 2017. By continuing to add more recently dated flight logs, the accuracy of the models could be improved with respect to relevancy by date.

5.3 Above Ground Level Error

The methods described in section 3 for altitude, as well as Appendix A, hold inherent error. The geospatial models, both ellipsoidal and geoidal, contain quantifiable error that could be reduced with the addition of flight log data containing laser altimeter generated AGL altitude.

6 Conclusion

This report included performance trends encapsulating about 30,000 flight logs from agricultural aircraft. The resulting distributions and standard deviations should be considered when designing the encounter geometry for testing DAA sensors and systems. Low flying agricultural aircraft present a challenge when attempting to account for the entirety of active aircraft in the current NAS. By analyzing a large data set of real-world flight logs, this research seeks to mitigate the risk associated with that challenge. Future work includes the maturation of probability models that seek to predict low altitude aviation traffic, as well as the integration of this research into the development of flight test planning for DAA encounter models.

Appendix A

The altitude values given were originally GPS altitudes. GPS altitude are a measure of the height of an ellipsoidal model of the Earth (see the red line labeled "Ellipsoid height, h"). In order to draw a meaningful comparison of the average altitudes for low flying agricultural aircraft, a better model of the Earth's surface is needed. For this reason, the GPS altitudes must be converted to an altitude Above Ground Level (AGL). In order to do this, there are two additional models that are required. The first is a terrain elevation model. For our analysis we used the GMTED 2010 7.5 arcsecond mean elevation data. This data sets gives an orthometric ground elevation (see the yellow line labeled "Orthometric height, H). The second model that is required is a dataset that gives the difference between a geoidal model of the earth and the ellipsoidal model (see the blue line labeled "Geoid height, N"). For this data, the GEOID 12B model was used. In order to calculate altitude AGL, the ellipsoidal GPS altitudes must be converted to orthometric altitudes. This is accomplished by referencing the Geoid12B model and subtracting the difference. Once the GPS altitudes are converted to orthometric altitudes, calculating approximate altitude AGL is done simply by subtracting the ground elevation from the elevation of the aircraft.

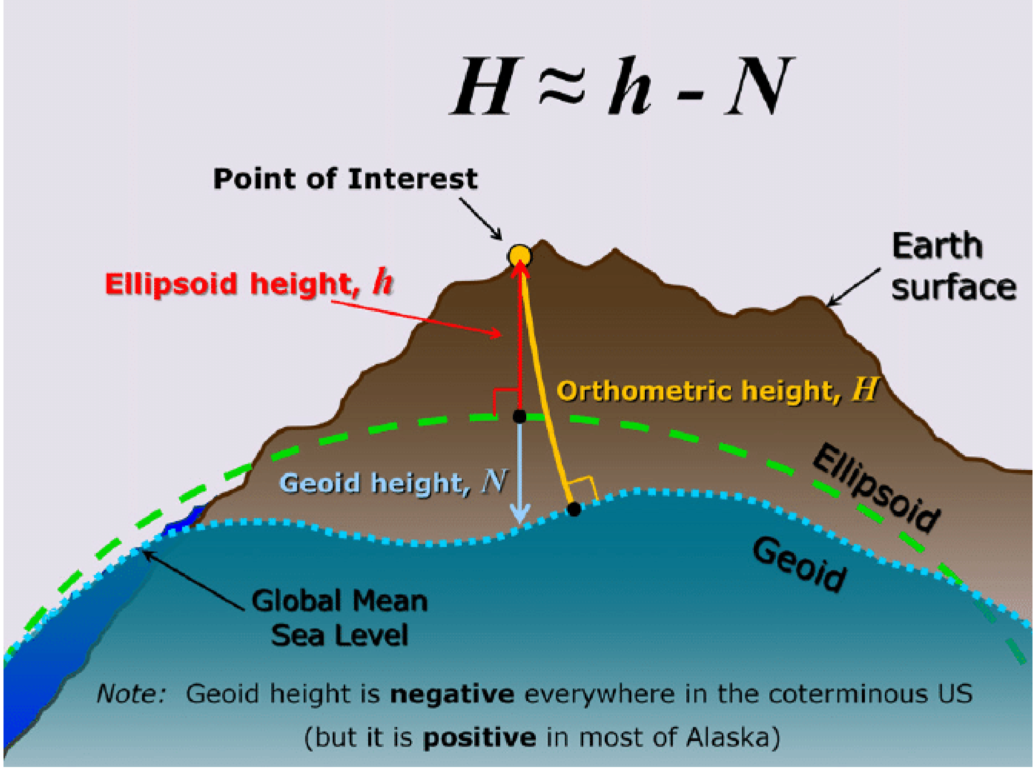


Figure 15. Conversion between orthometric and models of earth.

Appendix B

The variety of pilot habits when initializing flight log data as well as the inconsistency in file size required the differentiation and exclusion of certain files. See below figures for the different types of files during the preprocessing of data. The following two graphs represent the acceptable file type.

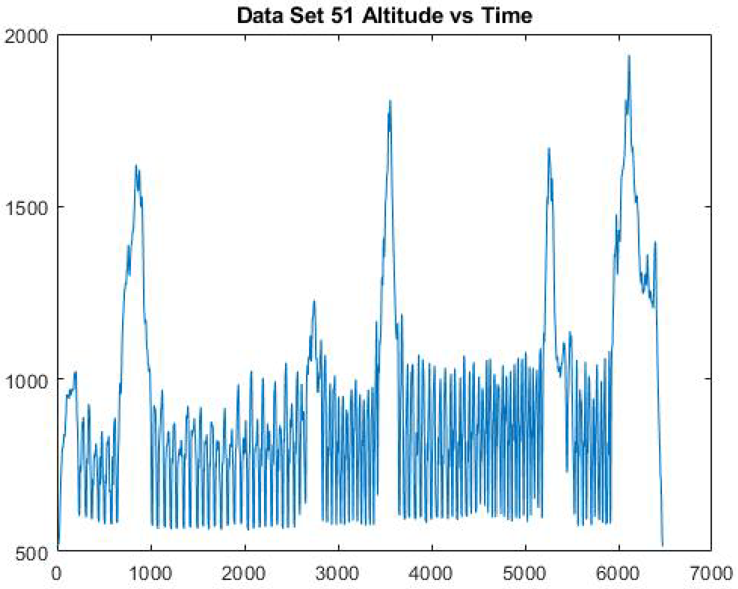


Figure 17. Complete File with Multiple Spray Runs and a Possible Landing.

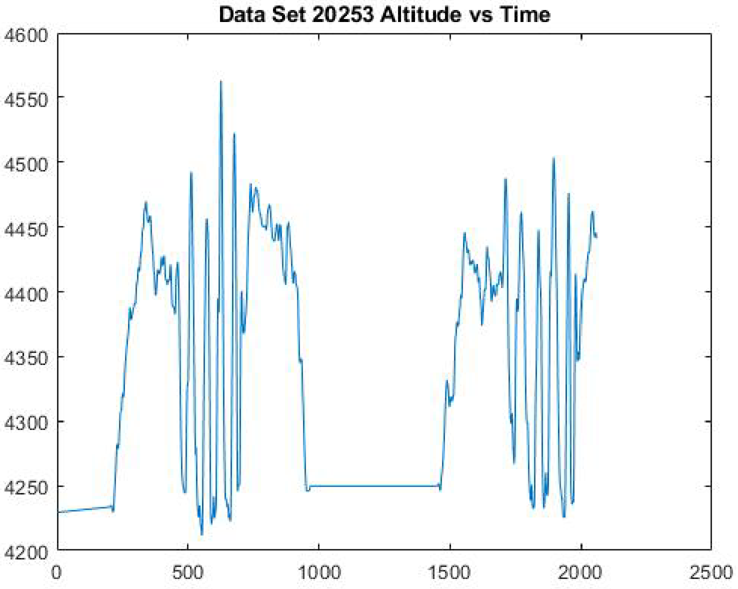


Figure 18. File with Travel to a Field and a Possible Intermittent Landing.

The above figures show clear variance in altitude over time. In Figure 17, a nearly complete file with multiple spray runs and possibly multiple unique fields is represented by the plot. In Figure 18, there exists at least two different sets of spray runs. The uniqueness of these fields is indeterminable by the graph but could still add to the data set. The following graphs represent unusable or unreliable data due to its short length or unreadable nature.

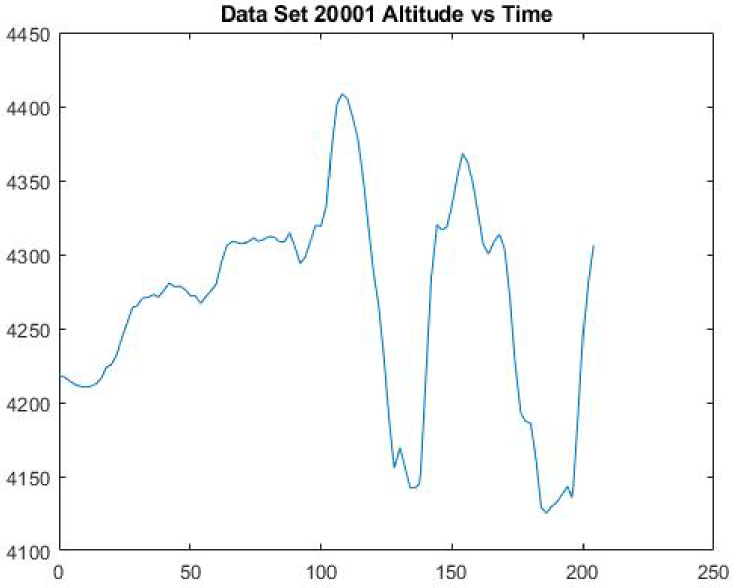


Figure 19. Possible Spray Run but not useful file.

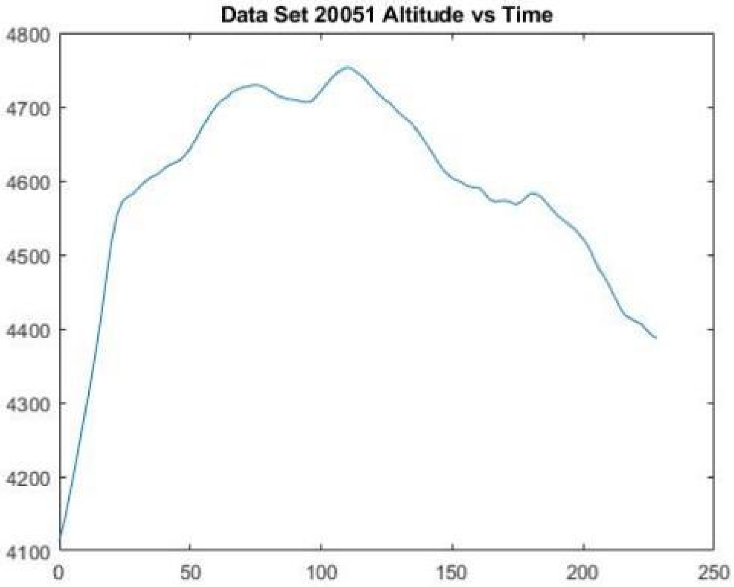


Figure 20. Possible cruise to a field but indeterminable.

Appendix C

Angles Relative to Earth's Surface

Method 1.1

This method goes through the entire dataset and loads each file's latitude, longitude, altitude, spray, and time properties into appropriately labeled vectors. It then goes through the spray matrix and takes every value of spray that is not 0 and sets it to 1. At this point delta vectors are created for the spray, altitude, latitude, and longitude vectors. The ground distance covered is estimated assuming a spherical earth using the formulas:

$$Ground\ Distance = \sqrt{Latitude\ Distance^2 + Longitude\ Distance^2}$$

$$Latitude\ Distance = \Delta latitude \cdot 365.223 \cdot 10^3$$

$$Longitude\ Distance = \ \Delta longitude \cdot 131480 \cdot 10^3 \cdot \frac{\cos\,(initialLatitude)}{360}$$

Once ground distance is calculated then the calculation for the descent angle begins. This is done by taking the average change in altitude and ground distance covered over a static window beginning 10 points before the spray turned on to 5 points before it turned on. With these two values the formula to calculate descent angle is:

$$\theta = \tan^{-1}\left(\frac{Average\ Change\ in\ Altitude}{Average\ Ground\ Distance}\right)$$

The ascent angle is calculated using the exact same formula except the window is calculated 5 points after the spray cuts off to 10 points after the spray cuts off. These values are then averaged for every file and the average ascent and descent angle for each file is stored in an array. Finally, each file's average ascent and descent are averaged to give the average climb angle and descent angle for the data set respectively. Each file is weighted equally.

Method 1.2

This method attempts to move away from a static data collection window but is otherwise identical to method 1.1. For the descent collection window, a function is called that takes in the current index of the file as well as the delta altitude and delta spray vectors and outputs the indices for the collection window. It starts at the current point that was fed to it by the parent function and then iterates backwards until it either hits another spray event or comes to two consecutive points where the change is altitude less than or equal to a predefined limit. Once it finds the second point with a change in altitude less than or equal to that limit it sets the stop index to the first value that was less than or equal to the predefined limit. The function then continues to iterate backwards through the delta altitude vector until it either goes 50 iterations backwards, hits the first component of the vector, or finds two consecutive elements that are greater than or equal to the predefined limit. Once the function finds two consecutive values that are greater than or equal to the predefined limit it sets the start index to the last value that was less than or equal to the predefined limit. If the function runs into another spray event, goes 50 iterations backwards, or hits the first element of the delta altitude matrix before setting the stop window, it will return a 0-start index and 0-stop index. If the function runs into a spray event, goes 50 iterations backwards, or hits the first element of the delta altitude matrix after setting the stop index it will set the start index to the current index value. The ascent window is calculated in a similar way. Instead of iterating backwards it goes forwards a max of 50 iterations. It first calculates the start index by finding two consecutive values that are greater than or equal to the predefined limit. Once it finds the second point that is greater than the limit it sets the first point as the start index. It then calculates the stop index by finding two points that are less than the limit. When it finds two consecutive points that are less than the limit it sets the last point that was greater than the limit as the stop index. Similarly to the descent window, if the function runs into another spray event, goes 50 iterations forward, or hits the last element of the delta alt matrix before it sets the start window the function will return a 0-start index and a 0-stop index. If the function runs into a spray event, goes 50 iterations forward, or hits the last element of the delta altitude vector after it has set the start index it will set the stop index to the current index value. An example dataset is shown below, demonstrating what the data collection window would be with a 5 ft/iteration limit (negative for descent, positive for ascent). A 1 in the delta spray column indicates the spray valve opening, a -1 indicates that the spray valve is closing. The orange boxes indicate the start of a spray run, the yellow boxes indicate the end of the spray run. The green boxes indicate the descent window that would be returned (start index 4, stop index 8). The blue boxes indicate the ascent window that would be returned (start index 21, stop index 26).

Table 9. Example file with descent in green and ascent in blue. The cutoffs for the algorithm defined in orange and yellow.

Delta Spray	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Delta Altitude	-1	1	-2	-5	-6	-4	-5	-6	-2	-1	-1	0	1	2	3	2	1	3	5	4	5	6	9	4	5	6	4	3	-1	1

Method 1.3

The idea for taking the turn rate of each flight log consists of evaluating the turn angles that each pilot takes for every turn between spraying runs. To achieve this, the periods of interest are defined as the times between the end of the previous spray and the beginning of the next spray. After searching for all changes in spray state, cleanup is performed to eliminate flight logs that do not have any transitions in spray state or cases that would not amount to significant data due to their length, like a pilot switching the spray system off and immediately back on. Then the turn rate is calculating by taking the latitude/longitude coordinates for every point, spacing out points as needed, and determining the curvature of the curve using the current point and the next 2 points in the curve. Next, this curvature is converted into a degree value and a turn rate is taken by dividing the degree values by the time between points to normalize the curves by position. As this value is an instantaneous value, the average is determined by taking the mean across all turns, and an average per flight log is taken by sub-setting each flight log as its own value. Examples of the turn rates can be seen in Figure 21.

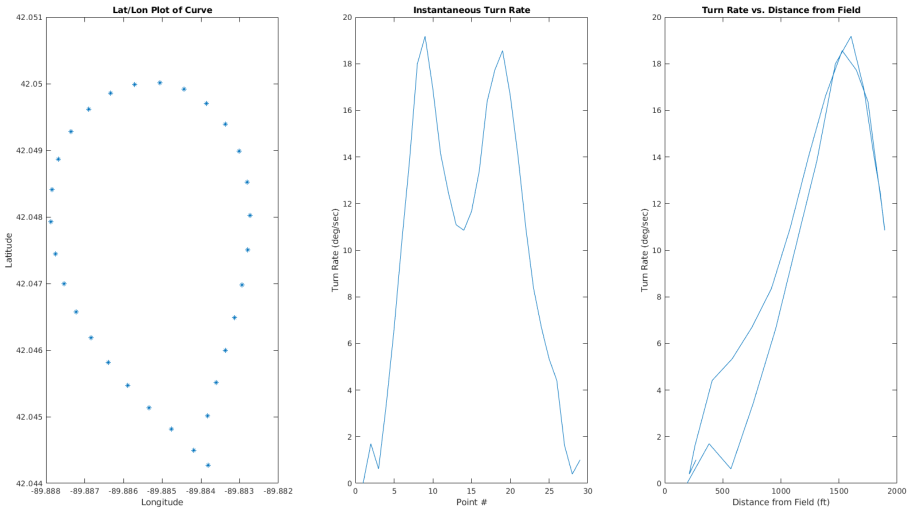


Figure 21. Turn rate between spray runs.

Spray Speed and Altitude

Method 2.1

This method goes through the entire dataset and takes the average speed while spraying for each file and assembles those averages into an indexed array. The output result is an average of this new array. Values that are above 770 mph (roughly the speed of sound), infinite, and NaN are disregarded. Each file is weighted equally.

Method 3.1

This method takes in each file’s latitude, longitude, GPS altitude, and spray vectors. For each file it goes in and creates an altitude above ground vector. It creates this by taking the GPS altitude, which is an altitude based on an ellipsoidal earth model, subtracting the difference between the ellipsoidal model and the geoidal model EGM2008 (to better approximate MSL altitude), and then subtracts the ground height in MSL provided by the mean elevation 7.5 arcsec model from the GMTED2010 database for every point. Once this altitude above ground vector is created, the altitude for every point during a spray run is calculated and the average is taken for each file. All files are weighted equally. Infinite and NaN values are discarded but this method does not filter negative or abnormally large values.

Method 3.2

This method begins by finding each file that begins while the plane is taking off. A file was determined to be a “takeoff file” if the initial speed was 50 mph or below and the change in altitude across the first 3 points of the file was less than 5 ft. Then this method takes all the takeoff files and subtracts their initial altitude from the entirety of the altitude vector to approximate altitude above ground. Then the average altitude while spraying is calculated by taking the mean of the new altitude vector at every point while the aircraft is spraying. Infinite and NaN values are discarded but this method does not filter negative or abnormally large values.

Method 3.3

Method 3.3 is similar to method 3.1 but uses the Geoid12B geopotential model instead of the EGM2008 model. Values are only accepted above -500 and below 500.

Cruise Speed and Altitude

Methods 4.1 and 5.1

The first method utilized the assumption that there were 3 possible times for the aircraft to be considered in cruise or "travelling" between fields. These were between takeoff and the first job, from one job to another, and from the final job to landing. This method began by defining what would be considered a job. The first step in this process was to look at the data's delta spray vector and cluster the spray runs into jobs based on the amount of time between spray runs. This initial cluster would iterate through the spray vector until it found another spray run or exceeded a limit. This limit was defined initially by a static 100 steps forward, but as the cluster got larger this was averaged to dynamically define clusters. Once the cluster reached 10 spray runs, the static 100 steps that was used to begin iterating through was deleted from the average and the program continued to iterate forward until it found another spray run or passed the limit defined by the following:

$$lim = mean(distance\ between\ runs) + 3 \cdot \sigma_{mean\ dist}$$

Once it eventually exceeded the limit, it set the start index and stop index for the job and then began looking ahead in the file for the next possible job. This initial clustering intentionally overpredicted the number of jobs but made sure that they were distinct enough to begin clustering the jobs using a geo-boundary method. This method defined polygons using the maximum and minimum latitudes as well as the maximum and minimum longitudes within each spray window as their vertices. It then calculated an estimated field center for each spray run by taking the average latitude and longitude over the initial job window. Example results are shown in Figure 22.

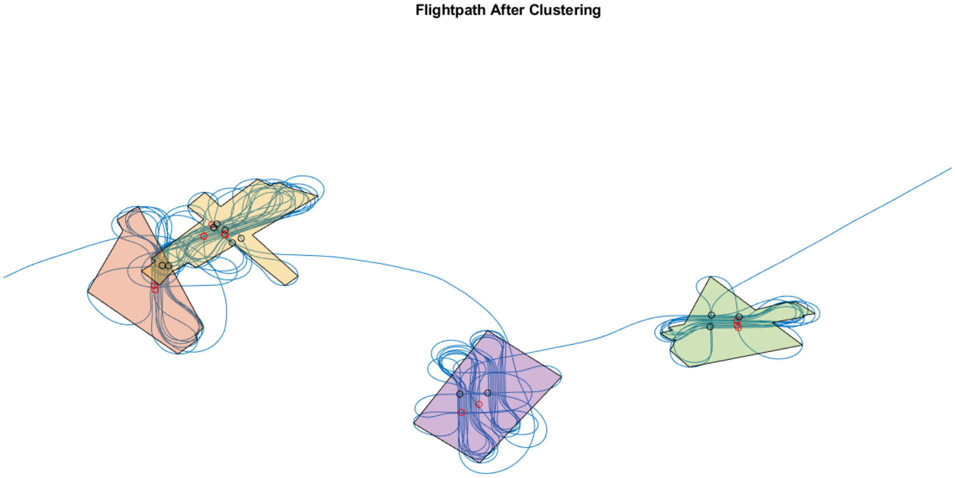


Figure 22. Geographically clustered polygons representing likely fields (2-D top down view).

The program then took each polygon and looked at both the preceding and following job windows to determine if the polygon contained either of their estimated field centers. If it did and there was less than 15 minutes between job windows, the program would meld the current polygon with that of the field center that it contained. Once the program had gone through all the original polygons it outputs the corresponding job windows to the newly melded polygons. Once the data set's jobs were defined, the program then looked at the beginning of the file and the end of the file to determine if the data contained a takeoff or landing. If it contained a takeoff it calculated a takeoff window from the beginning of the file to the first job. If the file contained a landing it calculated a landing window from the last spray run of the last job to the end of the file. It then went on to define windows between jobs. Once the windows were calculated, the processes from method's 2.1 and 3.3 were used to calculate the average cruise speeds and cruise altitudes respectively. The spray speed came out within reason, but the travel altitude came out to be an average of less than 100 ft across the entire dataset. Upon further investigation, this was caused by the false assumption that the data files would only contain takeoffs and landings at the very beginning or very end of the file respectively.

Methods 4.2 and 5.2

This method defines cruise as any time that the aircraft is travelling above the mean job altitude plus 2 standard deviations of the job altitude across a file and the aircraft is not currently in a job. To begin, the function calculates the file's job windows using the process defined in method 4.1 and 5.1. It then goes through these windows and calculates the average GPS altitude across all jobs within a file. Once it has these values calculated it sets a minimum cruise altitude limit equal to the average altitude plus 2 standard deviations of the job altitude average across the file. It then goes through each spray, latitude, longitude, and altitude vectors and deletes the portions of them that fall within a job window. Once it has these reduced vectors it then goes through them once again and removes each value where the corresponding altitude at that point is below the minimum cruise altitude limit. At this time, the average cruise speed across the file is calculated and stored in an array. Then the GPS altitude is converted to altitude above ground level using the methodology defined in method 3.3 and the average cruise altitude across the file is calculated and stored in an array. These arrays are then cleaned up (any infinite, NaN, or 0 values are removed) and the average of each is taken to generate the average cruise speed and average cruise altitude above ground for the entire dataset.