A Frequency Analysis Approach for Categorizing Air Traffic Behavior

Michael Drew*
University of California, Santa Cruz at NASA Ames Research Center,
Moffett Field, CA 94035-0001, USA

and

Kapil Sheth†
NASA Ames Research Center, Moffett Field, CA 94035-0001, USA

A method of analyzing National Air Space (NAS) air traffic that uses the Discrete Fourier Transform (DFT) is presented. The DFT is used to transform time domain traffic count data into the frequency domain where the sources of traffic in air spaces can be identified and characterized more easily. It is shown in simulation that individual traffic flows within Air Route Traffic Control Centers can be distinguished by their periodicity in the DFT plot. Next, three Traffic Management Initiatives (playbook rerouting, metered flows, and Ground Delay Programs) are implemented in simulations and their signature effects on the traffic are identified using the DFT. Finally, historical flight data is studied and the DFT is applied to sector traffic count data. It is found that in many cases, variations in traffic due to rerouting and convective weather disturbances are better highlighted in the frequency domain than in the original time domain data. Initial results of the DFT show it has potential as a tool for measuring and/or predicting NAS behavior for daily tactical planning and control purposes.

I. Introduction

T

raditional methods of characterizing air traffic in the National Air Space (NAS) have relied on measurements such as number of operations, flight counts, delay (total, average, and peak), airspace demand, etc. These data are useful for understanding and analyzing the performance of the national air transportation system, and are the basis for the majority of the work in the air traffic management field. Several methods of modeling, predicting, and optimizing the performance of the NAS using these metrics have been proposed. For example, Sridhar et al. construct a linear time-varying model of aggregate traffic flow using flight counts as the state variable. An autoregressive model for predicting sector demand is presented by Chen and Sridhar in Ref. that uses historical data and takes convective weather conditions into account. There has been less work in applying frequency analysis techniques to air traffic data, though related spectral analysis techniques that operate on the stochastic properties of the traffic have been successfully used to model and predict both air and automotive traffic. In Ref. for example, a spectral analysis on airport performance is conducted based on histograms of arrival rates. The authors found that the approach was useful for graphically distinguishing the operational differences between airports. In Ref. a spectral analysis technique of using historical traffic flow covariance matrix modal functions is used to forecast automotive traffic. Results were found to be comparable to other prediction techniques based on time series models.

In this paper, the authors have begun to investigate the use of the Discrete Fourier Transform (DFT) on airspace traffic count data. The DFT is used to characterize various NAS airspace regions, and to detect disturbances in the those regions that are not as apparent in the time domain. Nominal traffic in several Air Route Traffic Control Centers (ARTCCs) is characterized by applying the DFT to simulated traffic data. Simulations of various delay-causing Traffic Management Initiatives (TMIs) such as playbook routes,
metering, and Ground Delay Programs are also studied in the frequency domain using the DFT. Finally, the DFT is used to search for those TMI frequency domain signatures in the historical data.

This paper is organized as follows: In Section II the mathematical background to the DFT is presented, and its method of application is described. The results of the analysis are shown in Section III which is divided into three parts. The first part presents a characterization of some ARTCCs using the DFT on simulated data. The second part demonstrates the use of the DFT for characterizing three TMs, and the third part shows the results of applying the DFT to historical flight data. Conclusions and future work are presented in Section IV.

II. Background

A. The Discrete Fourier Transform

This section begins with a brief overview of the Fourier transform, which transforms a signal from the time domain into the frequency domain. It is based on the idea that a signal in time can be decomposed into a sum of sines and cosines over an infinite range of frequencies. The resulting plot of the Fourier transform is the amplitude of the sinusoids at those frequencies. As an example, consider the signal shown in Fig. 1a, which is made up of 3 sinusoids of different frequency and amplitude. If a sample of this signal is long enough to be assumed infinite, the plot of the Fourier transform yields 3 distinct spikes shown in Fig. 1b at 10, 100, and 250 Hz with amplitudes 4, 2, and 3 respectively. Thus, the signal has been transformed into the frequency domain and no information has been lost. In fact, given the signal of Fig. 1b, the time series signal can be reconstructed using what is known as the Fourier transform synthesis equation.

\[
x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{i2\pi kn/N},
\]

where the complex \(X[k]\) terms are the Fourier series coefficients. Equation 1 is known as the DFT synthesis equation, or the inverse DFT equation. One way to interpret this result is that any \(N\)-length discrete sampled time series can be thought of as the sum of \(N/2+1\) sines and cosines at \(N/2+1\) discrete frequencies ranging from 0 to \(F_s/2\), where \(F_s\) is the sampling frequency. The amplitudes of the sines and cosines are given by

Figure 1: Signal \(x(t) = 4\sin(2\pi(10t)) + 2\sin(2\pi(100t)) + 3\sin(2\pi(250t))\) in time and frequency domains.

The original Fourier transform was developed and prescribed for continuous time signals. Since traffic data, like most time series data in the modern world, are sampled at discrete intervals, a version of the Fourier transform known as the Discrete Fourier Transform (DFT) is used. A sequence of \(N\) time series data samples, \(x[0], x[1], ..., x[N-1]\) can be represented as

\[
x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{i2\pi kn/N},
\]
the $X[k]$ terms, which come from the DFT analysis equation:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-i2\pi kn/N}. \tag{2}$$

As with the continuous time Fourier transform, once the signal has been transformed into the frequency domain, (using Eq. (2)), it can be completely reconstructed using Eq. (1) despite the DFT having only a finite $N/2 + 1$ number of frequency points. In practice, the Fast Fourier Transform is used to efficiently calculate equation (2), and in this paper MATLAB’s ‘fft’ function is used to generate the results. Fourier transforms (specifically the DFT) show up in countless applications in science and engineering. Much has been written about their theory and application. See Refs. 6 and 7 for more details.

B. Window Functions

Implicit in the mathematics of the DFT is the assumption that the time series signal $x[n]$ is infinite and periodic—that is, the signal repeats itself from negative infinity to positive infinity. This has to be the case because sines and cosines are defined as extending from negative to positive infinity, thus $x[n]$ in (1) is not finite. Furthermore, $x[n]$ must be periodic since an infinite number of sinusoids are required to synthesize an aperiodic signal, and the DFT uses only a finite number of frequencies.

Since real-world signals are not infinite, and most are not periodic, the DFT is calculated with the assumption that values of $x[n]$ for $n < 0$ and $n > N − 1$ are all zero. This process is known as windowing because it is analogous to filtering out all samples of an imaginary infinite and periodic signal except for those in $0 \leq n \leq N − 1$. When calculating the DFT of a finite aperiodic signal using the Fast Fourier Transform (FFT) algorithm, this simple type of window function, known as a rectangular window, is implicitly applied. Unfortunately, for reasons beyond the scope of this paper, windowing a signal produces artificial anomalies in the DFT plot. There are alternative methods of windowing, however, that trade off some negative characteristics of those anomalies. The proper choice of a windowing function is critical and varies depending on application.

In order to determine the appropriate windowing function for this type of data, a single air traffic simulation was generated that contained a known frequency signature in an Air Route Traffic Control Center (ARTCC, or Center). (This specific simulation will be referenced later in the paper with more details.) Using one day’s recorded center counts from this simulation, and applying four popular windowing functions to the data prior to computing the DFT, produces the results shown in Fig. 2. Here, and in the remainder of this paper, time domain is plotted against period instead of frequency because it is more intuitive in this application. Recall that period $T$ is simply the inverse of frequency $f$, and since the count data is sampled at 1-minute intervals, the resulting period is in minutes.

The simulation run in Fig. 2 is known to have periodic content at approximately 13.4 minutes, and all four plots indicate this. The smaller spikes to the left of 13.4 minutes are the harmonics of this fundamental periodic content. Periodic signals often produce a spike at their fundamental frequency $f_f$ along with smaller spikes at multiples of that frequency $2f_f, 3f_f, \text{etc.}$ (In terms of fundamental period, these spikes are shown at $T_f/2, T_f/3, \text{etc.}$) While the rectangular window results shown in Fig. 2a preserves the DFT amplitude the best, it also contains a lot of artificial noise—especially near the fundamental frequency spike at 13.4 minutes. Using this window, it might be difficult to notice important periodic events in the DFT plots. The Blackman-Nuttall window shown in Fig. 2d smooths out the most noise of all four windows, but it also severely reduces the amplitude of the fundamental frequency. The Hanning and Hamming windows shown in Figs. 2b and 2c are very similar in their formulations and results. Throughout the remainder of this paper, the Hamming windowing function will be used to condition the time series count data prior to calculating the DFT. This window function is one of the most commonly used windows in signal processing, and strikes an acceptable balance between noise reduction and amplitude attenuation for this research. For more information on window functions, see Refs. 6, 7, and 8.
Figure 2: A comparison of four window functions applied to the same simulated traffic count data. DFT amplitude versus period.

III. Results

The results of this analysis are divided into three parts. In the first subsection, the DFT is applied to selected centers in an effort to characterize their nominal traffic. In the second subsection, various traffic management initiatives (TMIs) are applied in simulation in order to identify their DFT signature. Finally, in the third subsection historical data are analyzed from the DFT perspective in an attempt to identify events in the NAS that would otherwise be overlooked from the raw time series data.

A. DFT for Characterizing Center Traffic

The DFT can be calculated for any element in the NAS that can produce a time series of air traffic counts. Examples of such time series include center or sector counts versus time, and airport departures versus time. This work begins by selecting 30 weekdays of traffic from 2012 with the lowest amount of NAS-wide delay in minutes. Using this data set, simulations of nominal traffic are run using the Future ATM Concepts Evaluation Tool (FACET)\textsuperscript{\textregistered}. Herein, nominal traffic indicates that only the filed flight plan is used to simulate the behavior of each flight. This is as opposed to simulations where additional controller actions have been simulated, or historical flight data, which is updated according to actual radar-track-recorded aircraft positions.

In Fig. 3a, the air traffic counts in Cleveland Center (ZOB) for each of the 30 days are plotted for 24 hours, sampled in 1-minute intervals. The blue line shows the mean of these counts. Treating each of the
30 days of count data as a time series signal, the DFT is computed (after the Hamming window of Fig. 2c is applied) and the results are shown in Fig. 3b. Again, the mean of these DFT results are shown in blue. Fig. 3b shows that, as expected, the dominant periodic content of the counts data occurs at 1440 minutes. This is the 24-hour cyclical behavior of traffic, and the DFT of any daily traffic data signal would exhibit this characteristic.

Figure 3: Simulated nominal traffic Cleveland Center counts and DFT for 30 selected days. Mean counts and DFT shown in blue.

More interesting information may be found by looking at the higher frequency (periodicity of less than 240 minutes) events that exist closer to the origin of Fig. 3b. For the remainder of this paper, the DFT plots will focus on this higher frequency range because it is expected that disruptions in traffic flow will be exhibited here. Also, while the periodic 24-hour cyclical nature of traffic is known, traffic events of lower periodicity are not as well understood or cataloged.

Of the 30 days represented in Fig. 3, the center count data of December 6, 2012 is closest to the mean counts (that is, the sum of the absolute error from the mean counts is the least), and thus the most similar to a “typical” low-delay delay. In Fig. 4 the center counts and DFT plots are shown for this date in black. Note that there is a spike at approximately 110 minutes in the DFT plot (Fig. 4b), indicating that a significant event exists in the data with that periodicity. Suspecting that this might be caused by flights originating from and heading towards Detroit Metropolitan Wayne County Airport (DTW), one of the major airports supplying traffic to Cleveland Center, these flights were subtracted from the center counts, and the DFT was recomputed. The remaining center counts and DFT results are plotted in blue in Figs. 4a and 4b. The DTW flight counts and DFT results are shown in maroon. It is clear from Fig. 4b that the spike at 110 minutes is primarily due to the DTW flights.
These results are compelling because they demonstrate the potential value of the DFT for flight count data analysis. Without any prior knowledge of what distinguishes the blue time series line in Fig. 4a from the solid black line, all that can be easily surmised is that the flight volume has been reduced. However, if it is known that flights to and from DTW airport in Cleveland Center have a signature periodicity of approximately 110 minutes, this volume reduction could be identified as being caused by a disruption in DTW airport operations by viewing the data in the frequency domain. A reduction in the DFT plot at approximately 110 minutes would suggest the reduction in center-wide traffic volume as a reduction in DTW flights.

Generally, air traffic count data in a center (or sector) is the result of several other contributing entities of the NAS (e.g., specific airports, neighboring air space, specific jet routes, etc.). If those entities exhibit a periodic signature, an operational disruption in them would be hidden in the time series data, but made more obvious in the DFT plot (frequency domain). This is clearly demonstrated in the Cleveland center DFT plot, and several other centers exhibit similar characteristics. For example, most of the spikes in the DFT of Denver Center (ZDV) counts can be attributed to flights to and from Denver International Airport (DEN). This is shown in Fig. 5 where the DFT of only DEN flights in Denver Center for May 14, 2012 (shown in maroon) shadows the DFT plot of total Denver Center counts. May 14 exhibited traffic most similar to the mean traffic in that center.

Unfortunately, it is not always clear how to explain the DFT plot in its entirety for all centers. This is partially due to the limited time series sample size in addition to the noise introduced by windowing as discussed above. It is suspected, however, that there are reliable explanations for the source of large spikes that appear in this range in center count DFT plots. Further analysis and interaction with subject matter experts may be needed to identify them. For example, in Fig. 6 when a similar analysis is done with Atlanta Center (ZTL) and Atlanta International Airport (ATL) for January 4, 2012, the major DFT spikes of the total center data (in black) cannot be attributed to ATL traffic (in maroon) despite ATL being a major source of traffic in that center. Again, the traffic of January 4 was most similar to the mean traffic in Atlanta Center.
B. Catalog of Initiatives

In times of convective weather, or traffic congestion of any sort, traffic management initiatives (TMIs) are implemented to reroute, meter, or delay aircraft. In this subsection, examples of some TMIs are implemented in simulation to observe and catalog their effects in the frequency domain via the DFT.

1. Playbook Rerouting

One of the common TMIs used by controllers to reroute traffic around convective weather is known as the severe weather avoidance playbook CAN₁_East. This involves rerouting east-bound flights heading toward the northeast airports (e.g., New York Metropolitan, Washington, Philadelphia, and Boston area) north into Minneapolis Center (ZMP) and over the Great Lakes region. The first high-altitude sector in ZMP to be affected by the increase in traffic is ZMP20. Thus, choosing the high volume date of July 6, 2012, and implementing the CAN₁_East playbook in simulation in FACET from 16:00 - 24:00 UTC, the traffic counts and DFT results for that sector are compared against the baseline nominal traffic data in Fig. 7. As expected, Fig. 7a shows that traffic volume increases substantially with the additional playbook rerouted aircraft. In the frequency domain, this additional stream of aircraft is identified by an increase in the amplitude of the DFT plot between 30 and 50 minutes of period. Though it is not yet understood why this playbook route increases the periodicity of the sector counts within this range, other sectors through which this playbook stream traverses exhibit similar behavior.

2. Miles in Trail Metering

If the CAN₁_East playbook-affected stream of aircraft is metered at the Rapid City waypoint (RAP), which is upstream of ZMP center, the metered playbook stream of traffic can be detected in the DFT plot. In Fig. 8b it is shown that without any metering, the traffic in ZMP11 shows an increase in amplitude between 35 and 55 minutes of period (similar to ZMP20). As the miles-in-trail (MIT) of separation is increased on this stream, the DFT plot approaches that of the baseline DFT curve. This is because as the playbook...
traffic stream is spread out in distance, the traffic density of the stream is reduced, and thus the increase in single sector counts due to this stream is lessened. In other words, the traffic looks more like the original baseline traffic—both in the time and frequency domains. The curve for 50 MIT does, however, show a telltale spike at approximately 6.7 minutes. This corresponds to the distance in time between flights traveling at approximately 450 knots. (The spike at 3.35 minutes for the 25 MIT simulation is harder to detect within the noise.) In fact, the plots shown in Fig. 2 were generated from the DFT of total ZMP20 sector counts with 100 MIT applied to the CAN_East flow, producing a spike at approximately 13.4 minutes.

![Figure 8: Sector ZMP11 Counts and DFT for July 6, 2012, 16:00 - 24:00 UTC.](image)

### 3. Ground Delay Programs

Spikes in the center count DFT plots can often be attributed to traffic associated with a major airport in the center. With this in mind, simulations were carried out to determine if Ground Delay Programs (GDPs), which involve delaying flights on the ground at their origins to limit the arrival rate at an airport, could be detected using the DFT. Four simulations were carried out involving Denver Center (ZDV) and Denver International Airport (DEN) using traffic data from May 14, 2012. The nominal arrival rate for DEN is 120 aircraft per hour. Denver Center counts are recorded from 12:00 - 20:00 UTC when DEN is at the nominal arrival rate, and then for 3 additional simulations where this rate has been significantly reduced. Figure 9 shows both the time series and DFT results for these simulations. Note that because the time series data is shorter (8 hours instead of 24) the nominal DFT plot takes a different shape than it does in Fig. 5. This is because the amount of data is reduced, so there is less resolution in the DFT plot. However, there is still a significant spike shown between 90 and 180 minutes of period. The simulated GDPs are not in full effect until 17:00 UTC, so the time series center count data of Fig. 9a is nearly identical from 12:00 - 17:00 UTC. Nevertheless, the GDP effects are contrasted nicely in Fig. 9b, with the DEN signature spike being progressively reduced as the arrival rate at DEN is reduced.

To clarify, arrival rates (measured in flights per hour) and traffic periodicity plotted in DFT plots (measured in minutes) are not the same thing. In other words, the nominal arrival rate of 120 flights per hour is not related to the spike in Fig. 5 being near 120 minutes. Rather, the reduction in arrivals at DEN has reduced the signature of DEN traffic within Denver Center, which happens to exhibit a periodicity of between 90 and 180 minutes. Also, it should be mentioned that a reduction in volume as exhibited in the time series data of Fig. 9a does not necessarily correlate to a reduction in DFT magnitude. Recall that the DFT is displaying the amplitude of periodic events that make up a time series signal. All things being equal, a mere reduction in mean traffic counts would produce an identical DFT plot at all periodicities except at the infinite-length period (0th frequency). In this work, all traffic count data has been mean-shifted prior to calculating the DFT in order to eliminate this effect.
C. Identifying Historical NAS Events

So far, the DFT has been effective at distinguishing characteristic traffic flows within the NAS as well as the effects of playbook, metering, and GDP TMIIs implemented in simulation. Unlike simulation results, however, historical traffic data includes the effects of controller and airline actions as they respond to traffic demands, weather constraints, and other tactical changes that are not represented by the original filed flight plans used in the simulated results. Thus, it becomes more difficult to identify specific traffic flows and TMI actions in the presence of all the other day-to-day operational variances. Nevertheless, with some investigation, specific events in the NAS data record can be highlighted by the DFT analysis.

From operational record data, the CAN1_East playbook was implemented on July 26, 2012 at 20:15 UTC and extending into the following day. In Fig. 11, the DFT of the historical counts in sector ZMP11 from 16:00 - 24:00 UTC are compared to that of three other days chosen from the list of low-delay dates surrounding July 26. This is done to reduce the effects of seasonal traffic variance. The CAN1_East playbook caused an increase in amplitude for the July 26 results (shown in green) at nearly the same period as the simulated CAN1_East playbook results shown in Fig. 8b. It is not known exactly what TMI events caused the nearby spikes shown for the July 6 (blue) and August 2 (maroon) results, but historic radar data shows that there was significant convective weather in the region, and traffic in Sector 11 was disrupted. June 28 (black), by contrast, was clear of any convective weather within the region for this time span.

On August 16, 2012, the Vulcan (VUZ) playbook, which routes flights from the west to northeastern destinations southward through Atalanta Center (ZTL), was put into effect from 15:00 UTC into the following day. These flights pass through sector ZTL23, and all flights destined to Newark, LaGuardia, JFK, and Teterboro airports were metered to 35 miles-in-trail. Again, in Fig. 11, the DFT results of this date are compared to those of three nearby days from the low-delay list of dates. As clearly highlighted by the green line, which is lower than the other dates through most of the plot’s domain, the effect of this metered flow through this sector is to reduce the amount of high frequency (low period) content of traffic. In the time

Figure 9: Denver Center Counts with varying ground delay programs for Denver International Airport, May 14, 2012, 12:00 - 20:00 UTC.

Figure 10: DFT of sector counts for ZMP11 for selected dates 16:00 - 24:00 UTC.
domain plot of Fig. 12 the volume of counts for this day is increased as expected. Time series data for ZMP11 on July 26 is not shown but is similar to Fig. 12 in that any trend that would identify the presence of a playbook stream is difficult to observe.

Although the three other dates were chosen to be low-delay days from a NAS-wide perspective, it is obvious from Fig. 11 that something is substantially different about August 7, 2012. Looking into historic NEXRAD radar data, it was discovered that there were severe thunderstorms throughout the region. In response to the weather, historic TMI records for this time period show that there were significant metering actions throughout Atlanta Center, and traffic in Sector 23 was disrupted, resulting in more high frequency (low period) events in the counts data. Note that this effect cannot be easily determined by the time domain counts data shown in Fig. 12. The DFT is shown to highlight these differences in traffic more effectively than the raw time series data.

Figure 11: DFT of sector counts for ZTL23 for selected dates 16:00 - 24:00 UTC.

Figure 12: Sector counts for ZTL23 for selected dates 16:00 - 24:00 UTC.

IV. Conclusions and Future Work

In this paper, a method of analyzing traffic data within the NAS airspace was developed and applied to simulated and historical data. It was shown that a frequency domain analysis provides an alternative method of analyzing and understanding the data. Traffic within control centers and sectors can be characterized by their DFT plots to a level of detail that is not possible using time domain plots. It was found that applying the DFT to simulated nominal traffic counts within a control center allowed specific sources of the traffic to be identified by their periodic signatures. Often, when those sources are disrupted (as with a GDP, for instance, or any other TMI) the disruption can be detected at the center-wide level in the frequency domain, but not in the time domain. As presented here, the effects of playbook rerouting, metering, and GDPs could all be detected using the DFT method on simulated data. Despite those TMI signatures being less obvious and coherent in the historical DFT analysis, they could still be detected in many cases. In fact, while looking for the signature of the Vulcan playbook in historic data, the detrimental effects of thunderstorms and the

American Institute of Aeronautics and Astronautics
resulting TMI actions were noticed in the DFT results of a different day. These results would have been difficult to detect by time series data alone. Thus, the DFT may be an effective tool for looking through historical data and checking if certain initiatives were in place, or for detecting significant weather events and their impacts on traffic.

It has been shown that delay-inducing events like TMIs produce frequency domain signatures within airspace regions of the NAS, and it is hypothesized that such signatures, to the extent they can be detected in historical data, may be useful for analyzing NAS behavior. Eventually, such methods may be used to predict NAS behavior for daily real-time planning and control purposes by detecting traffic events more precisely. To succeed in doing so, more work needs to be done to identify the frequency domain signatures of specific traffic sources in centers and sectors, along with other initiatives like Airspace Flow Programs (AFPs). Each airspace is different, and varies in its frequency domain response to specific TMI actions. Also, because windowing and sample size play a major role in the DFT resolution and accuracy, more work needs to be done to determine how to best apply the technique—especially with historical data where controller actions tend to have smaller duration and flight plans are not followed as closely as they are in simulation.

References

5. MATLAB, version 8.0.0.783 (R2012b), The MathWorks Inc., Natick, Massachusetts, 2012.