Semi-Stochastic Aircraft Mobility Modelling for Aeronautical Networks: An Australian Case-Study Based on Real Flight Data

Jiankang Zhang, *SM, IEEE*, Luping Xiang, Dong Liu, Jingjing Cui, Soon Xin Ng, *SM, IEEE*, Robert G. Maunder, *SM, IEEE*, Thomas Graeupl, Uwe Carsten-Fiebig, Lajos Hanzo, *FIEEE*

Abstract—Terrestrial Internet access is gradually becoming the norm across the globe. However, there is a growing demand for Internet access of passenger airplanes. Hence, it is essential to develop aeronautical networks above the clouds. Therefore the conception of an aircraft mobility model is one of the prerequisite for aeronautical network design and optimization. However, there is a paucity of realistic aircraft mobility models capable of generating large-scale flight data. To fill this knowledge-gap, we develop a semi-stochastic aircraft mobility model based on large-scale real historical Australian flights acquired both on June 29th, 2018 and December 25th, 2018, which represent the busiest day and the quietest day of 2018, respectively. The semi-stochastic aircraft mobility model is capable of generating an arbitrary number of flights, which can emulate the specific features of aircraft mobility. The semi-stochastic aircraft mobility model was then analysed and validated both by the physical layer performance and network layer performance in the case study of Australian aeronautical networks, demonstrating that it is capable of reflecting the statistical characteristics of the real historical flights.

I. INTRODUCTION

Although every successive wireless generation has improved the attainable throughput, their focus has been on terrestrial global coverage of specific areas, where the high teletraffic holds the promise of lucrative revenues for the service providers. Specifically, the terrestrial 5G wireless system is capable of providing 20 Gigabits-per-second (Gbps) peak data rates and 100+ Megabits-per-second (Mbps) average data rates. By contrast, there is limited Internet access on passenger airplanes, which is also quite costly.

Hence, the concept of aeronautical *ad-hoc* networks (AANETs) has emerged for the provision Internet-Above-the-Clouds [1]. Passenger airplanes can also be harnessed as mobile base stations [2] for filling the large coverage holes in sparsely populated areas, where it is challenging to deploy and maintain the wireless infrastructure, as exemplified in Fig. 1 for Australia. However, the design, analysis, and optimization of aeronautical networks critically hinge on the topology and mobility of the aircraft. Hence, it is vital to develop a realistic and reliable aircraft mobility model for the design, analysis and optimization of aeronautical networks.

As one of the essential prerequisites, the mobility modeling of roaming users and/or vehicles is a topic that has received

J. Zhang is with Department of Computing & Informatics, Bournemouth University, BH12 5BB, U.K. (E-mail: jzhang3@bournemouth.ac.uk).

L. Xiang, D. Liu, J. Cui, S. X. Ng, R. Maunder and L. Hanzo are with School of Electronics and Computer Science, University of Southampton, SO17 1BJ, U.K. (E-mails: {l.xiang,d.liu,jingj.cui}@soton.ac.uk, {sxn,rm,lh}@ecs.soton.ac.uk).

Thomas Graeupl and Uwe Carsten-Fiebig are with DLR, Germany

Fig. 1. The 3G/4G/5G network coverages by Telstra, which are obtained from https://www.nperf.com/en/map/AU/-/2445.Telstra/signal/ on February 21st, 2021. The network coverages by Optus and Vodafone are similar or smaller than those covered by Telstra.

much attention due to its pivotal role in network characterization and network planning [3]. Hence, a methodical study of the associated mobility models is imperative in support of network development and optimization both in terms of the routing/scheduling protocols in the network layer as well as the associated signal processing algorithms of the physical layer.

Diverse mobility models have been conceived for mobile *ad hoc* networks (MANETs) [4], vehicular *ad hoc* networks (VANETs) [5], [6], and Flying *ad hoc* Networks (FANETs) [7] in order to evaluate their network performance and to design suitable networking protocols. The classical random direction model and random waypoint model have been investigated both in MANETs [8] and VANETs [9] as well as FANETs [10], [11], since they are capable of capturing the mobility characteristics of both users, as well as of manned and unmanned aerial vehicles (UAVs). However, although these *ad hoc* networks share the intrinsic feature agile mobility, they also differ quite significantly in a range of specific aspects, such as their velocity, direction pattern, network size, and geographical coverage area as well as topology, which require bespoke mobility modeling. Thus, the mobility models developed for different *ad hoc* networks have to capture the unique features of nodes.

Ad hoc networking between passenger airplanes was initially conceived by scientists at the German Aerospace Center (DLR) [12] under the terminology of airborne mesh networking. The terminology of 'airborne networks' has also been widely used in the context of UAVs and manned cargo planes. Hence we will commence from characterizing the mobility developed for *airborne networks* constituted both by unmanned and manned aircraft (passenger airplanes).

Numerous contributions have been devoted to developing

mobility models for airborne networks [13]–[17]. In order to improve their mobility-aware routing protocol design, Tiwari *et al.* [13] defined a trajectory-based mobility model for airborne networks constituted by wide-body aircraft based on the flight plans of all the aircraft participating in the airborne network backbone. However, this mobility model highly relies on the pre-defined flight plan recorded, hence it is not scalable for network expansion. Furthermore, it only investigated four nodes in a area upto $2500m \times 2500m$, which is unsuitable for aeronautical networks constituted by commercial passenger airplanes. As a further development, the authors of [14] proposed a spiral line based mobility model for improving the smoothness of the synthetic trajectory. However, the backbone nodes in [14] are UAVs rather than the wide-body aircraft of [13]. A 3-dimensional (3-D) Gauss-Markov mobility model was proposed in [18] for modelling the mobility of aircraft flying at different altitudes, which was also referred to as a multi-tier mobility model [19]. Rohrer *et al.* [15] developed a 3-dimensional Gauss-Markov mobility model. As a further advance, Li *et al.* [20] improved the smoothness of a Markov-like mobility model by incorporating three extra states, namely the pre-decelerate, turn, and post-accelerate states. A semi-random circular movement mobility model was developed in [16] for simulating UAVs circling around a specific centre in order to collect information in a particular target area. Wan *et al.* [21] proposed a 2-D smooth turn based random mobility model (2D-ST-RMM) by capturing the correlation of the acceleration of airborne vehicles across both the temporal and spatial coordinates, which is capable of reflecting the characteristics of making smooth turns of different turn radii, as well as capturing the tendency of aircraft travelling straight and making turns of large radii. Five years later, Wan *et al.* [22] developed a pair of realistic 3D-ST-RMMs for capturing the diverse mobility patterns of fixed-wing aircraft, which relied on coupling stochastic forces with the physical laws that govern 3-D aerial maneuvers. In their 3D-ST-RMMs, the model features were also introduced for determining the movement of aerial vehicles, when they approach simulation boundaries. Sharma *et al.* [17] conceived a mixed mobility model for UAVs by combining the random waypoint based and the uniform mobility models for characterizing the movements of a UAV in vertical and horizontal directions, respectively. The above mobility models may be classified as stochastic mobility model, which is capable of reflecting the randomness well owning to their intrinsically random mechanism. However, the capability of capturing randomness is attained at the cost of degrading their capability of reproducing realistic mobility patterns and the topology in emulating the flight-path of commercial passenger airplanes, which are closely related to the airport distribution over a given geographical area and time period. Furthermore, none of the existing models has been validated by real flight data.

By contrast, Graeupl created the Framework for Aeronautical Communication System evaluation 2 (FACTS2) [23] to support the development of new aeronautical data links for air traffic guidance in Europe. As discussed in Graeupl's paper, the FACTS2 air traffic mobility model relies on the analysis of the years 2007 and 2008. Two reference days of average air traffic were analysed for hourly aircraft generation rates for each pair of airports in the database. The aircraft generation rates were then extrapolated into the future by applying growth factors published by the European organisation for the safety of air navigation referred to as EUROCONTROL. Flights were then simulated to fly along certain routes between the identified airport pairs with take-off times modelled as a stationary Poisson process according to the hourly aircraft generation rates. The results have been shown by Graeupl to model European air traffic quite closely [24].

The mobility models of [14], [16], [17], [20]–[22] have been developed for UAVs, which exhibit distinctly different features from those of passenger and cargo aircraft. Hereinafter, the term '*aircraft*' will refer to passenger aircraft. However, at the time of writing, there is a paucity of mobility model developed for commercial passenger airplanes. Nevertheless, Graeupl *et al.* [23], [24] targeted both passenger and cargo aircraft under the so-called Instrument Flight (IFR) rules. Although their model could also be used for air-to-air communications, it has only been used so far for the simulation of air-to-ground communication in support of air traffic guidance in part of the European airspace, rather than for investigating AANETs linking aircraft to airport and for investigating the network performance.

In Table I we boldly and explicitly compare the main contributions of [13], [15], [18], [23], [24] to ours in this paper, because [13], [15], [18], [23], [24] have the most similar features to those of the passenger aircraft targeted by us. By observing Table I, we can see that the existing mobility models have never been validated by real flight data in the open literature. Last but no least, there is a paucity of realistic aircraft mobility models capable of capturing the node distribution over a given geographic area and over a given time period. Against this background, we develop a semi-stochastic aircraft mobility model for the Australian airspace that is capable of generating an arbitrary number of flights capturing the flight distribution over a period of 24 hours, reflecting both the real topology and the mobility pattern. Explicitly, our main contributions can be summarized as follows:

(1) For the first time, we characterize the mobility features of passenger aircraft based on large-scale real historical flight data gleaned from flights in Australia, which includes the top-5 airlines' flights on two representative dates.

(2) We develop a semi-stochastic mobility model for passenger aircraft, which is capable of generating an arbitrary number of flights over Australia. In contrast to fully stochastic mobility models that cannot accurately capture the features of real flight mobility, our semi-stochastic aircraft mobility model inherits the capability of capturing near-realistic aircraft mobility, topology as well as distribution over a given geographic area and time period, while exhibiting the required element of randomness.

(3) We proposed a single-source-to-multiple-destination routing optimization scheme based on Dijkstras algorithm, which is capable of finding the *best routing path* either in terms of the end-to-end delay quantified by the number of hops, or in terms of the end-to-end throughput from a target aircraft to multiple ground stations (GSs) as destinations.

(4) For the first time, we investigated the Australian AANET based on large-scale real flight data both in the physical and network layers. Furthermore, we demonstrated that our semi-stochastic aircraft mobility model is capable of

TABLE I
COMPARISON THE MAIN CONTRIBUTION OF MOBILITY MODEL FOR AIRBORNE NETWORKS.

Analysis/Validation	[13]	[18]	[15]	[23], [24]	Ours
Real flight data				A & V	A & V
Topology					A & V
Distribution				A & V	A & V
Connection ratio					A & V
Area SE					A & V
Throughput	A			A & V [1]	A & V
Latency	A				A & V
Link life time					A & V
Packet delivery			A		
Node	WBA	Aircraft	Aircraft	IFR Aircraft	Passenger Aircraft
Comments				European A2G ATC networks	

A: Analysis   V: Validation   A & V: Analysis and validation   Blank : No related work

1 For the air-to-ground networks developed for air traffic control. For the LDACS network, the simulation results have also been validated by measurements.

emulating the real flights over the Australian airspace. Hence, the model-based and real data based networks exhibit similar performance.

(5) Finally, our semi-stochastic methodology is also applicable to generating flights in other airspaces, such as the dense airspaces over populated areas represented by Europe, the United States (US) and China. Furthermore, our methodology is also suitable for modelling the airspace over unpopulated areas represented by the North-Atlantic (NA) region. Hence some statistical characteristics of the dense European and NA airspace are also presented.

The rest of this paper is organized as follows. Section II presents the methodology of developing our semi-stochastic aircraft mobility model. In Section III, we present a distance-based adaptive coding and modulation scheme specifically designed for quantifying the link quality between a pair of aircraft. The single-source-to-multiple-destination routing optimization scheme based on Dijkstras algorithm is discussed in Section IV. Section V is devoted to the analysis and validation of the semi-stochastic aircraft mobility model both in the physical and in the network layer. The applicability of our semi-stochastic aircraft mobility model to other airspaces is discussed in Section VI. Finally, in Section VII, we conclude and briefly discuss our future research ideas.

II. THE METHODOLOGY OF AIRCRAFT MOBILITY MODEL

In this section, we present our methodology of generating aircraft traffic based on large-scale historical flight data exemplified by the Australian scenario. Nevertheless, the methodology presented below is also applicable to other scenarios. The international airlines only have routes flying from/to the international airports in Australia to/from other countries' international airports. Our methodology of generating near-realistic aircraft mobility is detailed as follows:

1) *Identify the flight designators.*
   - *Identify the airlines having most flights*. Intuitively, each airline typically schedules a single international flight each day connecting a pair of international airports with one of them located in Australia, whilst the other one being in other countries/continents. Hence, the number of international flights of a single airline is typically less than that of a domestic airline's flights.
   - *Identify the flight designators of the top-5 Australian domestic airlines*. By checking the flight schedules

Number	2018−06−29	374	358	186	42	47
of flights	2018−12−25	243	287	211	61	0
Airlines	QANTAS	Virgin australia	Jetstar	tigerair	reX. Regional Express

Fig. 2. The top-5 Australian airlines and their number of flights scheduled on June 29th, 2018 and December 25th, 2018.

of the airlines' official website and the Australian airports' website, the top-5 airlines in terms of the number of flights scheduled on each day are Quantas, Jetstar, Virgin Australia, Tigerair and Rex (Regional Express). Explicitly, their number of flights scheduled on June 29th 2018 and December 25th 2018 are shown in Fig. 2, where Christmas day of 2018 represents the *quietest day* having the least flights in 2018, whilst the date of June 29th 2018 represents the *busiest day* having the most flights in 2018.

2) *Download historical flights on the quietest day and on the busiest day.*
   The historical flight data can be downloaded either manually or automatically from Flightradar24 https://www.flightradar24.com/. There are diverse sources available for obtaining historical flight data, where commercial license is required, such as Flightaware, Flightradar24, Flight Tracker - OAG, Automatic Dependent Surveillance-Broadcast (ADSB) -exchange, Aviation Edge, FlightStats, Flightfinder and FlightView. All of them provide live tracking functions for flights. However, historical flight data are mainly provided by their so-called data service at a fixed charge per query. Flightradar24 allows users to customize their preferred historical flight data, hence we opted for it as our source of the top-5 Australian domestic airlines' flight data on June 29th 2018 and December 25th 2018, respectively.

3) *Harmonise the raw flight data.*
   The historical flight data downloaded from Flightradar24 is in raw format, which is based on a comma-separated value (CSV) file containing floating point numbers, text and special characters. It is not directly usable for analysis. The user has to first read and change the hybrid format raw data to floating point number format, which can be used for our following analysis. Furthermore, the time intervals of raw data for each flight data are random, which is unsuitable for analysis and validation. Hence, we harmonised the flight data by employing a unified time interval of $\Delta_t = 10$ seconds for each flight data. Explicitly, given two time-data pairs (t_0, y_0) and (t_1, y_1),

the linear interpolation [25] of (t, y) is given by

$$\frac{y - y_0}{t - t_0} = \frac{y_1 - y_0}{t_1 - t_0}. \tag{1}$$

Furthermore, the data value y for a particular time instant t can be calculated as

$$y = y_0 + (t - t_0)\frac{y_1 - y_0}{t_1 - t_0}. \tag{2}$$

With the aid of the linear interpolation operation of Eq. (2), each data entry of altitude, latitude, longitude, and speed can be interpolated at time intervals of $\Delta_t = 10$ s. To exemplify our interpolation, we present the flight data of Jetstar's JQ574 taking off from Melbourne and heading to Brisbane on June 29th, 2018 as a specific example. Explicitly, Fig. 3(a) - Fig. 3(d) present our comparison between the interpolated flight data of Jetstar's JQ574 and the real raw flight data of JQ574 in terms of altitude, latitude, longitude and speed, respectively. There are several typical phases of flight behaviours, such as landing and taxiing, as well as climbing and en-route, plus the potential holding phase encountered, when the aircraft approaches the airport but has no clearance to land. Our semi-stochastic aircraft mobility model exploits a large scale dataset of real historical flights. Explicitly, each flight data contains the entire flight path trajectory, based-on the latitude, longitude, altitude, and speed seen in Fig. 3 all of the above-mentioned flight-phases may be readily recognized.

The timestamp on the x-axis is Unix Epoch time (UET), which can be converted into a date and universal time coordinated (UTC) format as follows:

$$\begin{aligned} \text{UTC time} = &\Big(\big(\left(T_{UET}/60\right)/60\right)/24\Big) \\ &+ \text{DATE}\left(1970, 1, 1\right), \end{aligned} \tag{3}$$

where T_{UET} is a timestamp in UET format.

4) *Archive the harmonised real flight data into a library.*
The harmonised real flight data are entered into our library, which are accessible as benchmarked flight data for generating artificial flight data as required.

5) *Characterise the distribution of take-off time.*
We assume the take-off time of each flight to be the instant of changing its speed from zero to non-zero. Hence the resultant take-off time may be slightly different from the scheduled take-off time published by the airport/airline, which is typically a little bit later than the published take-off time. Nevertheless, our method represents the '*real*' take-off time, which provides more accurate flight status and topology than simply relying on the published take-off time. Moreover, flight delays routinely occur in reality due to weather, Air Traffic Control (ATC) restrictions, bird flocks, knock-on effects, etc.

In possession of the extracted take-off times, the statistical characteristics can be analysed. Explicitly, the probability density function (PDF) of take-off times on June 29th 2018 and December 25th 2018 are depicted in Fig. 4(a) and Fig. 4(b), respectively. The corresponding cumulative distribution functions (CDFs) of the take-off times on June 29th 2018 and December 25th 2018

are depicted in Fig. 4(c) and Fig. 4(d), respectively. The histogram of take-off times was then approximated by classical distribution functions, such as the Weibull, Gamma, Normal, Poisson and Students-t distribution. Furthermore, we have also included the Kernel density estimation (Kernel distribution) as a benchmark. However, Kernel Density Estimation does not constitute a convenient parametric technique of estimating the PDF of a random variable. Hence we will not adopt it for generating the artificial take-off time, regardless of its goodness-of-fitting accuracy.

The accuracy of distribution fitting can be tested by a suite of test methods, such as the classic Chi-squared test, Kolmogorov-Smirnov test (KS-test), Hosmer-Lemeshow test, and Kuiper's test, etc. The Chi-squared test has been widely used, but its accuracy depends on having a sufficiently large sample size for the approximations to be valid. By contrast, the KS-test does not rely on the sample size for its inference to be valid and it does not depend on the specific shape of the CDF under test. Hence, we have adopted the KS-test for quantifying the accuracy of distribution fitting. Given the CDF $F(x)$ and the empirical distribution function F_n for n independent and identically distributed (i.i.d.) ordered observations, the KS-test statistic is defined as [26]

$$D_n = \sup_x |F_n(x) - F(x)|, \tag{4}$$

where $\sup_x$ is the supremum of the set of distances, while D_n is also known as the goodness-of-fit, which is typically used for evaluating the accuracy of a specific fitting operation. For the KS-test, a smaller value of D_n represents a better fit.

The goodness-of-fit values of the six hypothesis distributions investigated are summarized in Table II. Again, without considering the non-parametric Kernel density estimation, the Normal and the Weibull distributions have the best goodness-of-fit for the take-off times on June 29th 2018 and December 25th 2018, respectively. More specially, the mean and variance of Normal distribution are $\mu_J = 14.18$ and $\sigma_J = 4.97$ for fitting the flight data on June 29th, 2018, whilst the scale value and shape value of the Weibull distribution are $\phi_J = 15.80$ and $\psi_J = 3.16$ for fitting the flight data on June 29th, 2018. By contrast, the mean and variance of Normal distribution are $\mu_D = 13.43$ and $\sigma_D = 4.47$ for fitting the take-off time of flights on December 25th, 2018, whilst the scale value and shape value of the Weibull distribution are $\phi_D = 14.98$ and $\psi_D = 3.36$ for fitting the take-off time of flights on December 25th, 2018.

6) *Randomly generate take-off time according to the distribution acquired in Step-5.*
Given the distribution of take-off times acquired by comparing the goodness-of-fit D_n, we can now artificially generate a set of take-off times. In order to provide further insights concerning the artificially generated flight data, we consider both the best fitting and the second best fitting distributions, hence both the Normal and the Weibull distributions will be considered for generating artificial flight data in our investigations. As shown in Fig. 5(a) and Fig. 5(b), the take-off times generated by the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2021.3104118, IEEE Transactions on Vehicular Technology

5

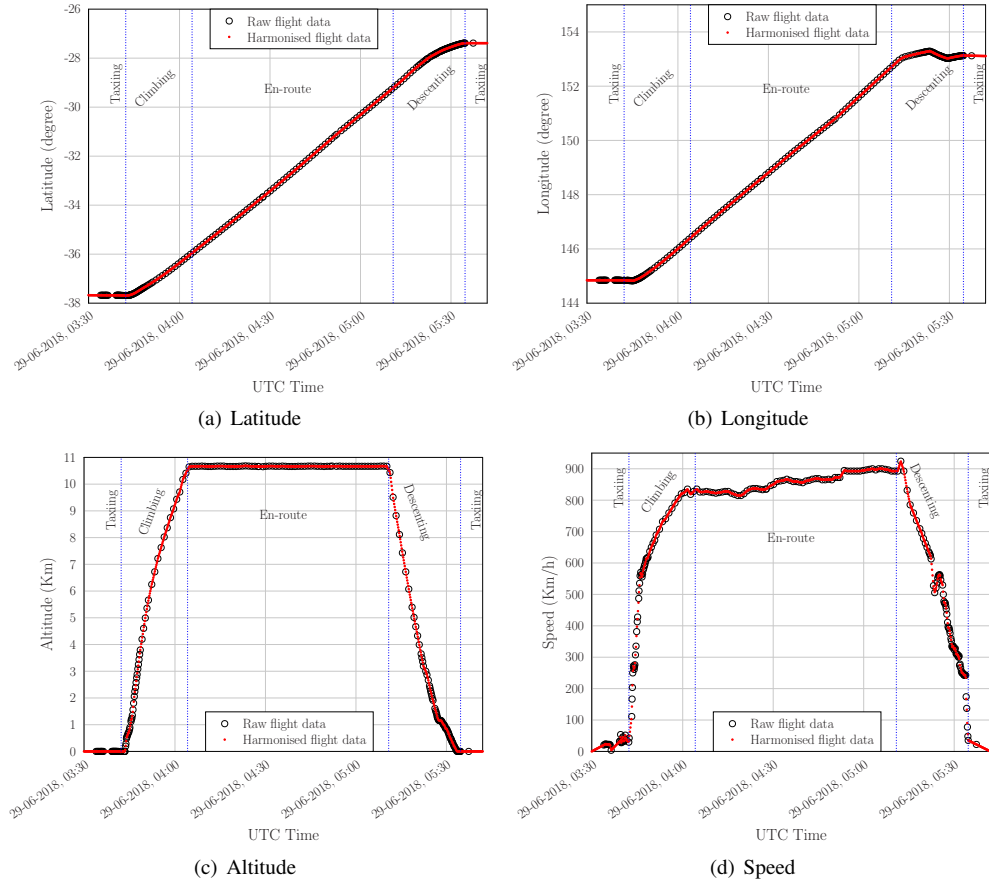(a) Latitude

(b) Longitude

(c) Altitude

(d) Speed

Fig. 3. Comparison between the harmonised and the real raw flight data for the Jetstar JQ574 taking off from Melbourne and heading to Brisbane on June 29th, 2018. The raw flight data is directly obtained from Flightradar24.

(a) Jun. 29th, 2018

(b) Dec. 25th, 2018

(c) Jun. 29th, 2018
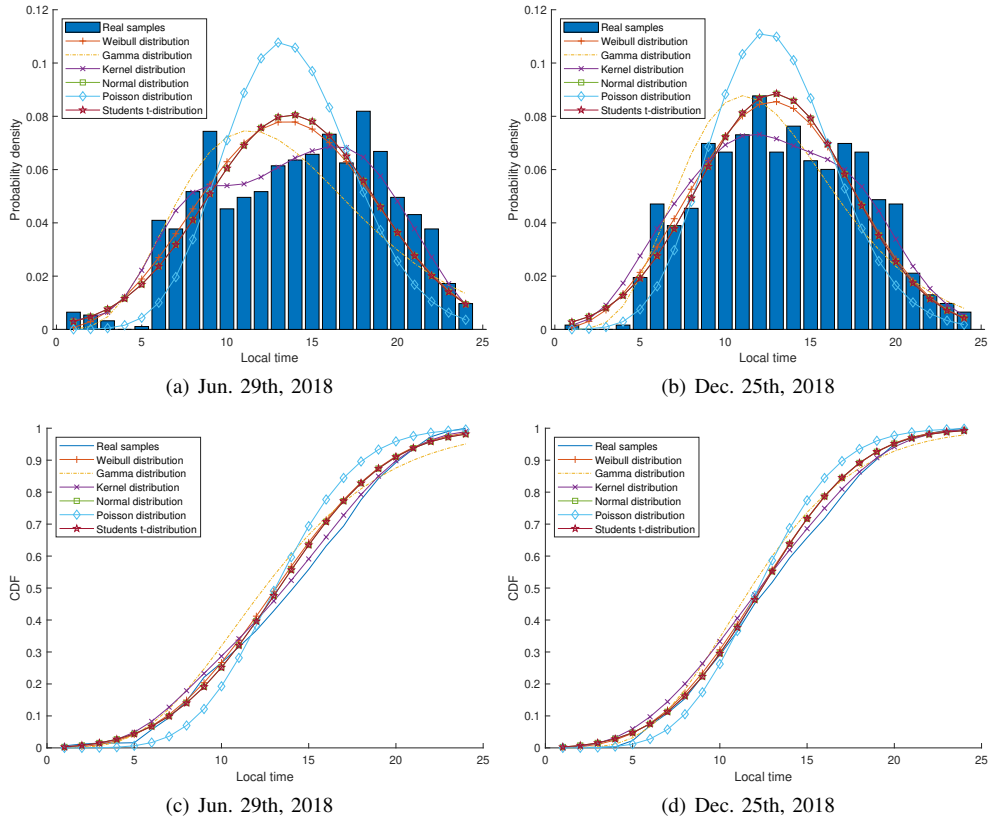
(d) Dec. 25th, 2018

Fig. 4. Characterizing the distribution of take-off time for the top-5 Australian airlines' flights. The total number of flights on June 29th, 2018 (December 25th, 2018) is given here by the sum of the top-5 Australian airlines' flights.

TABLE II
GOODNESS-OF-FIT D_n FOR DIFFERENT STATISTICAL MODELS OF THE DISTRIBUTION OF TAKE-OFF TIMES

Distributions	Weibull	Gamma	Kernel	Normal	Poisson	Student-t
June 29th,2018	0.0848	0.1131	0.0339	**0.0782**	0.1504	0.0783
December 25th, 2018	**0.0680**	0.0818	0.0460	0.0691	0.1266	0.0694

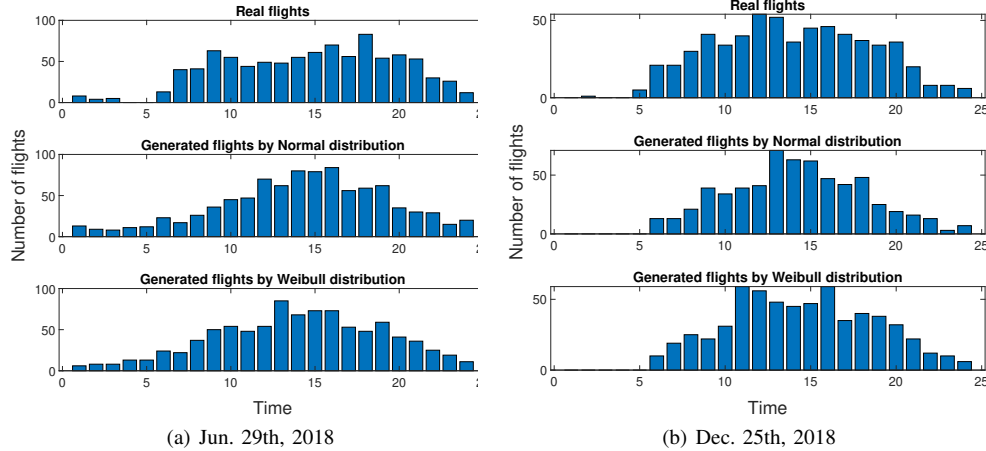(a) Jun. 29th, 2018

(b) Dec. 25th, 2018

Fig. 5. Comparing the artificially generated take-off times to the Normal and Weibull distributions

Normal and Weibull distribution are capable of accurately matching the take-off time of the real historical flights both on June 29th, 2018 and on December 25th, 2018.

7) *Shift the real historical take-off to "randomly generated take-off time"*. A set of randomly selected real historical flight data extracted from our database will be used as baseline flight data. The associated take-off time is first extracted by selecting the first non-zero speed value and its corresponding timestamp. Let us denote the take-off time of the randomly selected real historical flight data by T_r. Furthermore, given a randomly generated take-off time T_g obeying either the Normal or the Weibull distribution, the artificially generated flight data associated with the speed interval of v_g is given by

$$v_g = \begin{cases} [\mathbf{0}_{N_s}, v_{r,1}, v_{r,2}, \cdots, v_{r,N-N_s}], & \text{if } T_g > T_r \\ [v_{r,N_s+1}, v_{r,2}, \cdots, v_{r,N}, \mathbf{0}_{N_s}], & \text{if } T_g < T_r \\ v_r, & \text{if } T_g = T_r \end{cases} \quad (5)$$

where v_r is a randomly selected real historical flight data velocity, $v_{r,i}$ is the i-th element of v_r, N is the length of vector v_r, $\mathbf{0}_{N_s}$ is a row vector of length of N_s having 0 elements, which is calculated as:

$$N_s = \frac{\mathcal{F}(T_g - T_r)}{\Delta_t}. \quad (6)$$

Again, recalling the operation in Step-3, the time interval between data samples is $\Delta_t = 10$ s. In Eq. (6), $\mathcal{F}(\cdot)$ is the operation converting the time difference between T_g and T_r into seconds. Similarly, we can obtain the artificially generated flight data of latitude θ_g, longitude φ_g, and altitude L_g following the methodology of Eq. (5). In a nutshell, the flight trajectory of the artificially generated flight data $(\theta_g, \varphi_g, L_g, v_g)$ follows a same route as the historical flight data $(\theta_r, \varphi_r, L_r, v_r)$, but they have different takeoff time and landing time.

III. DISTANCE-BASED ADAPTIVE CODING AND MODULATION

In mesh networks of aircraft, the ground stations (GSs) are typically located at the airport, which allows the aircraft to directly communicate with Air traffic control for delivering vital control messages. Apart from being either the source or destination, the aircraft also provide relaying services during their landing/take-off, taxiing and holding patterns [1], [27], where these phases are associated with rather diverse channel characteristics, especially the air-to-ground (A2G) or air-to-air (A2A) links [28]. The A2G communications and A2A communications rely on the same channel model for their data transmission, but they have different maximum delay. In the following, we present our distance-based ACM regime used for both A2G and A2A communications, nevertheless, the A2G data transmission relying on distance-based ACM follows the same methodology.

Intuitively, there is a line-of-sight (LOS) path in A2A communications, where the power ratio between the LOS path and the diffuse components is given by [27]

$$K_{\text{Rice}} = \frac{\alpha^2}{\beta^2}, \quad (7)$$

where $\alpha \in \mathcal{R}$ is the amplitude of the LOS path and $\beta \in \mathcal{R}$ is the variance of the diffuse component having zero-mean quadrature components.

Hence, given the Rician factor K_{Rice}, we have [27], [29]

$$\alpha = \sqrt{\frac{K_{\text{Rice}}}{1 + K_{\text{Rice}}}}, \quad (8)$$

$$\beta = \sqrt{\frac{1}{1 + K_{\text{Rice}}}}. \quad (9)$$

In the aeronautical channel, the Doppler power spectrum and delay power spectrum is dependent on the specific phase of flight. The Doppler power spectral density function of the A2A aeronautical channel can be modelled by Clarke's [30] formula given by

$$p(f_D) = \begin{cases} \frac{1}{\pi f_{D_{\max}}\sqrt{1-(f_D/f_{D_{\max}})^2}}, & |f_D| < f_{D_{\max}}, \\ 0, & \text{else,} \end{cases} \quad (10)$$

where $f_{D_{\max}}$ is the maximum Doppler shift. In the worst case, the direction of the LOS path coincides with the heading of the aircraft, hence resulting in a carrier shift of $f_{D_{LOS}} = f_{D_{\max}}$ for

the LOS path, whereas the scattered components arrive from behind.

Furthermore, the pathloss of the A2A aeronautical channel can be modeled as [29]

$$L_{\text{path loss}} \, [\text{dB}] = -154.06 + 20 \log_{10}(f) + 20 \log_{10}(d), \quad (11)$$

where f [Hz] is the carrier frequency and d [m] is the distance between the transmit antenna (TA) and receive antenna (RA).

The small-scale fading is characterized by the diffuse components of the multi-path radio channels, which can be written as [31]

$$h_{\text{NLOS}}(\tau, t) = \lim_{N \to \infty} \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j\theta_n} \cdot e^{j2\pi f_{D_n} t} \cdot \delta(\tau - \tau_n), \quad (12)$$

where τ is the excess delay, t is the absolute time, N is the number of echos, θ_n represents the phase of the n-th echo, and f_{D_n} is the Doppler shift of the n-th echo.

By contrast, the line-of-sight (LOS) component is given by [31]

$$h_{\text{LOS}}(t) = e^{j2\pi f_{D_{\text{LOS}}} t} \cdot \delta(\tau - \tau_{\text{LOS}}), \quad (13)$$

where $f_{D_{\text{LOS}}}$ is the Doppler frequency of the LOS path, and we have $\tau_{\text{LOS}} = 0$ after time synchronization at the receiver side.

Hence, recalling Eq. (12) and Eq. (13), the time-domain channel impulse response (CIR) $h(\tau, t)$ can be formulated as

$$
\begin{aligned}
h(\tau, t) &= \alpha \cdot h_{\text{LOS}}(t) + \beta \cdot h_{\text{NLOS}}(t), \\
&= \alpha \cdot e^{j2\pi f_{D_{\text{LOS}}} t} \cdot \delta(\tau - \tau_{\text{LOS}}) \\
&+ \beta \cdot \lim_{N \to \infty} \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j\theta_n} \cdot e^{j2\pi f_{D_n} t} \cdot \delta(\tau - \tau_n). \quad (14)
\end{aligned}
$$

We employ the distance-based adaptive coding and modulation (ACM) scheme developed in [29], [32]. Explicitly, the distance-based ACM aeronautical communication system can switch its ACM mode based on the distance between the desired pair of communicating aircraft.

Given the position of a pair of aircraft, say the transmit aircraft a and the receive aircraft b, we can calculate their 3D separation according to Eq. (15) - Eq.(17). Explicitly, having the altitude, latitude and longitude $(L_a, \theta_a, \varphi_a)$ of aircraft a, its Cartesian coordinates $p_{x,a}, p_{y,a}$ and $p_{z,a}$ are defined by

$$p_{x,a} = (R_E + L_a) \cos(\theta_a) \cos(\varphi_a), \quad (15)$$
$$p_{y,a} = (R_E + L_a) \cos(\theta_a) \sin(\varphi_a), \quad (16)$$
$$p_{z,a} = (R_E + L_a) \sin(\theta_a). \quad (17)$$

The 3D separation between aircraft a and aircraft b is calculated as

$$d_{a,b} = \sqrt{|p_{x,a} - p_{x,b}|^2 + |p_{y,a} - p_{y,b}|^2 + |p_{z,a} - p_{z,b}|^2}, \quad (18)$$

where $d_{a,b}$ is in [m]. In (4), $p_{x,a}, p_{y,a}$ and $p_{z,a}$ are the 3D Cartesian coordinates of aircraft a, while $p_{x,b}, p_{y,b}$ and $p_{z,b}$ are those of aircraft b.

Furthermore, the distance-based ACM scheme conceived for aeronautical communications using $K = 7$ modes is given in Table III, which can be designed according to our investigations in [29]. Note that the default system parameters in Table IV have been used in designing the distance-based

TABLE III
DISTANCE-BASED ADAPTIVE CODING AND MODULATION SCHEME FOR AERONAUTICAL COMMUNICATIONS.

Mode k	Mode color	Throughput (bps/Hz)	Threshold d_k (km)
0	None	< 0.459	> 740.8
1	Red	0.459	500
2	Orange	1.000	350
3	Yellow	1.322	200
4	Green	1.809	110
5	Blue	2.197	40
6	Magenta	2.747	25
7	Purple	3.197	5.56

ACM of Table III. The transmit aircraft, say aircraft a^*, selects an ACM mode to transmit its data according to the adaptive reconfiguration regime of

$$\text{If } d_k \leq d_{b^*}^{a^*} < d_{k-1} : \text{ choose mode } k, \quad (19)$$

where aircraft b^* is the receive aircraft, $k \in \{0, 1, 2, \cdots, K\}$, and we assume $d_0 = D_{\max}^{\text{A2A}}$. When $d_{b^*}^{a^*} \geq D_{\max}^{\text{A2A}}$, there exists no adequate communication link, since the two aircraft are beyond each others' communication range. Since the minimum flight safe separation must be obeyed, we do not consider the scenario of $d_{b^*}^{a^*} \leq D_{\min}$. Explicitly, $D_{\min} = 5.56$ km is used as the minimum safe separation distance, and $D_{\max}^{\text{A2A}} = 740.8$ km is the maximum communication distance having a non-zero throughput in A2A communication. When the distance exceeds $D_{\max}^{\text{A2A}}$, there is no adequate communication link.

The distance-based ACM designed in Table III is also suitable for A2G communications, but the maximum A2G communication distance is given by $D_{\max}^{\text{A2G}} = 370.4$ km as limited by the radio horizon [1]. Hence, only the orange, yellow, green, blue magenta and purple ACM modes are available for A2G communication.

IV. SINGLE-SOURCE-TO-MULTIPLE-DESTINATION ROUTING OPTIMIZATION BASED ON DIJKSTRAS SHORTEST PATH FIRST ALGORITHM

Given the multihop nature of airborne mesh networks, the packets have to be delivered through multiple wireless hops to arrive at their final destination, such as a GS for accessing the Internet. Hence, a seamless source to destination path must be established in an aeronautical *ad-hoc* network (AANET) to enable an aircraft to access the Internet [33], which is the task of routing protocols. However, having an appropriate address allocation is essential for facilitating reliable packet exchange between airborne mesh network nodes. Airborne mesh networks are self-configuring networks consisting of aircraft nodes interconnected by wireless links, where an efficient routing strategy has to be conceived for avoiding congestion and for achieving the maximum throughput per aircraft. Intensive efforts have been devoted to routing protocol design and routing optimization [34]–[38]. However, the choice of routing protocols critically hinges on the specific topology and mobility. The routing problem of the Australian aeronautical network - similar to other AANETs - typically has multiple destinations, because an aircraft may potentially direct its packets to any of the ground stations in order to access the Internet. Hence, the above-mentioned routing algorithms may be invoked for validating our semi-stochastic aircraft mobility model in terms of its network layer performance. However, further research efforts are required for optimizing the routing

algorithms for the multiple-destination Australian aeronautical network, which is set aside for our future research. To expound further Dijkstra's algorithm [39] is capable of finding the shortest distance or maximum end-to-end throughput to all other nodes from that node of origin. In our investigations, we will invoke Dijkstra's algorithm to find the shortest path from an aircraft to any GS as well as to find the best path maximizing the end-to-end throughput, which will be used both for characterizing our mobility model in terms of the number of hops imposed, the end-to-end throughput and the area spectral efficiency.

Dijkstra's algorithm [39] is a powerful search technique, which was specifically designed for finding the shortest path between a pair of nodes quantified in terms of the number of hops, each of which may of course have a different length in Km. When $A = 6$ aircraft and $B = 2$ ground stations are considered, we may construct the $C \times C$ throughput matrix W shown in Fig. 7, where $C = A + B$. Explicitly, the number in each grid position represents the throughput in bps/Hz conveyed between a pair of nodes using the ACM modes of Table III. Once this hop-count or throughput matrix has been constructed, Dijkstra's search algorithm can be used for finding the minimum number of hops from a certain aircraft to the GSs. It also allows us to determine the maximum achievable throughput for a certain aircraft at a complexity order of $\mathcal{O}(C^2)$ [39]. It is worth noting that the max throughput of a specific route is limited by that of the lowest-throughput hop.

In the following, we will further augment Dijkstra's algorithm by the quantitative example of Fig. 6. Explicitly, we illustrate both the links between different aircraft and the links between the aircraft as well as the ground station, relying on the corresponding throughput matrix shown in Fig. 7. For example, aircraft-5 is three hops away from GS-2 and in this route, it has a connection to aircraft 2 and the throughput of this link is 1.000 [bps/Hz], as seen in the throughput matrix of Fig. 7 and Fig. 6 as well as in the ACM mode of Table III. Similarly, the only aircraft connected to aircraft 2 in its route to GS-2 is aircraft 1, with the corresponding throughput of 2.197 [bps/Hz] stored in the throughput matrix, as seen in Fig. 6 and Fig. 7 as well as in the ACM mode of Table III. The throughput of the link between aircraft 1 and the GS-2 is also shown in Fig. 6 and again the throughput matrix of Fig. 7. Hence, then end-to-end throughput of the routing path 'aircraft-5 → aircraft-2 → aircraft-1 → GS-2' is 1.000 [bps/Hz], which is limited by the gold ACM mode of Table III. Alternatively, there is another routing path for aircraft-5 accessing GS-1, namely the routing path 'aircraft-5 → aircraft-4 → aircraft-3

Fig. 7. A stylized throughput matrix of six aircraft and two ground stations for a hypothetical example network. Explicitly, the number in each grid position represents the throughput in bps/Hz conveyed between a pair of nodes using the ACM modes of Table III.

	Aircraft–1	Aircraft–2	Aircraft–3	Aircraft–4	Aircraft–5	Aircraft–6	GS–1	GS–2
Aircraft–1		2.197	1.322					1.809
Aircraft–2	**2.197**				**1.000**			
Aircraft–3	1.322			1.000		0.459	1.809	
Aircraft–4			1.000		0.459			
Aircraft–5		1.000		0.459				
Aircraft–6			0.459					
GS–1			1.809					
GS–2	**1.809**							

Algorithm 1 Dijkstra's Max-throughput search algorithm

Input
  W, the starting aircraft index st, and the ground station index e;

Output
  the max throughput from aircraft index st to the ground station index e

Require:
  Initialise throughput vector C with the st column of W;
  Initialise $i = 0$;
  Initialise a flag n-length vector, ***visit***, to check whether the aircraft or ground station has been visited before;

1: **while** $i < n - 1$ **do**
2:    ***temp*** = **0**; //Initialise a n-legnth zero vector
3:    $j = 0$;
4:    **while** $j < n$ **do**
5:      **if** $visit[j] == 1$ **then**
6:        $temp[j] = C[j]$;
7:      **end if**
8:      $j = j + 1$;
9:    **end while**
10:    $(value, index) \leftarrow \max(\boldsymbol{temp})$; //Pick the neighbour aircraft that has the max capacity.
11:    $visit[index]$ = 0;
12:    $j = 0$;
13:    **while** $j < n$ **do**
14:      **if** $C[j] < \min(C[index], W[index, j])$ **then**
15:        $C[j] = \min(C[index], W[index, j])$; //Update the current max-throughput $C[j]$
16:      **end if**
17:      $j = j + 1$;
18:    **end while**
19:    $i = i + 1$;
20: **end while**
21: **return** $C[e]$;

→ GS-1'. However, its achievable throughput is limited by the red ACM mode between aircraft-5 and aircraft-4. Thus, the end-to-end throughput is 0.459 [bps/Hz], which will not be selected owing to its low throughput.

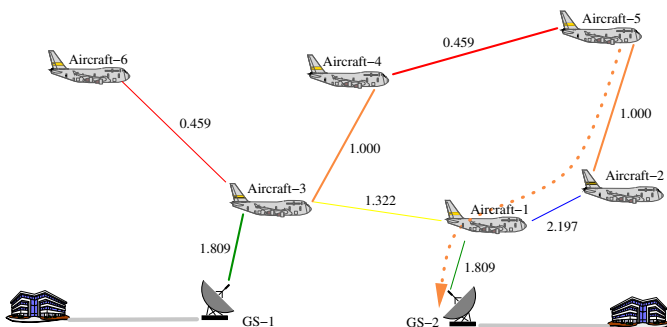More generally, Dijkstra's max-throughput search algorithm operates using the following steps:

Fig. 6. An example topology showing the links of the aircraft and ground station. The corresponding throughput matrix is shown in Fig. 7.

TABLE IV
PARAMETERS USED IN VALIDATING THE SEMI-STOCHASTIC AIRCRAFT MOBILITY

Scenario related parameters	Airspace	Australian airspace
	Top-5 domestic airlines	Quantas, Jetstar, Tigerair, Virgin Australia and Rex
	Number of GSs	15
	Representative dates considered	December 25th, 2018 June 29th, 2018
	Time period considered	$00:00 \sim 24:00$
	Total number of flights on December 25th, 2018	802
	Total number of flights on June 29th, 2018	1007
	Latitude	Determined by each aircraft
	Longitude	Determined by each aircraft
	Altitude	Determined by each aircraft
Physical layer parameters	Carrier frequency f_c	5 GHz
	Bandwidth B	6 MHz
	Number of CPs	32
	Number of subcarrier	512
	factor K_{Rice}	5 dB
	ACM	As detailed in Table III
	Minimum separation distance $D_{\min}$	5.56 km
	Maximum A2A communication distance $D_{\max}^{\text{A2A}}$	740.8 km
	Maximum A2G communication distance $D_{\max}^{\text{A2G}}$	370.4 km
Semi-stochastic mobility model based on Normal distribution	Total number of flights on June 29th, 2018	1007
	Mean μ_J on June 29th, 2018	14.18
	Variance σ_J on June 29th, 2018	4.97
	Total number of flights on December 25th, 2018	802
	Mean μ_J on December 25th, 2018	13.43
	Variance σ_J on December 25th, 2018	4.47
Semi-stochastic mobility model based on Weibull distribution	Total number of flights on June 29th, 2018	1007
	Scale value ϕ_J on June 29th, 2018	15.80
	Shape value ψ_J on June 29th, 2018	3.16
	Total number of flights on December 25th, 2018	802
	Scale value ϕ_J on December 25th, 2018	14.98
	Shape value ψ_J on December 25th, 2018	3.36

1) Initialise the max-throughput vector as $\boldsymbol{C} = \boldsymbol{W}_{st}$, where $\boldsymbol{C}$ is a column of $\boldsymbol{W}$.
2) Pick the source aircraft st and calculate the throughput of the link leading to the adjacent aircraft and to the ground stations.
3) Pick the next aircraft u_i having the max throughput, and if the capacity is higher upon visiting u_i, then update the current max-throughput C.
4) Repeat the adjacent aircraft throughput calculations until all the aircraft have been visited.

Dijkstra modified max-capacity search procedure is summarized in Algorithm 1, which is capable of achieving the maximum end-to-end capacity. Similarly, the minimum latency imposed by the end-to-end transmission can be acquired upon replacing the throughput by the delay quantified in terms of the number of hops. For example, when using a typical 5G-style transmit frame-structure, each node may add 10 ms of relaying delay upon additionally taking into account the node-distance in Km and the speed of light, the actual propagation delay in seconds may be readily found.

V. VALIDATING THE AIRCRAFT MOBILITY MODEL

In this section, we will analyse both the achievable physical layer link performance and the network layer performance in terms of the number of hops, latency and link life time in end-to-end transmission. Again, the dataset used for our analysis and validation is based on the top-5 domestic airlines' flights on June 29th, 2018 and December 25th, 2018, namely on Quantas, Jetstar, Tigerair, Virgin Australia and Rex (Regional Express). The aircraft mobility model developed will also be characterized in terms of its achievable link quality both in the physical and network layer.

The default system parameters used for our analysis and simulations are summarised in Table IV. Explicitly, this has a

bandwidth of $B = 6$ MHz and a carrier frequency of $f_c = 5$ GHz. Orthogonal frequency-division multiplexing (OFDM) is used and the transmit power is set to 1 watt per antennas. A typical Rician factor of $K_{\text{Rice}} = 5$ dB is considered for the aeronautical channel, as seen in Table. IV. Again, the default system parameters of Table IV have been used in designing the distance-based ACM of Table III.

In the following, '*Real*' in the legend represents the real historical flights. Furthermore, '*Normal*' represents the flights generated by our aircraft mobility model developed by relying on the Normal distribution for the take-off time, which represents 'the flights generated by the Normal distribution'. By contrast, the label '*Weibull*' represents the flights generated by our aircraft mobility model relying on the Weibull distribution for the take-off time, which represents 'the flights generated by the Weibull distribution'. The semi-stochastic aircraft mobility model developed in this paper is capable of generating an arbitrary number of aircraft scheduled on a date required for the Australian aeronautical networks. However, in order to validate its capability of capturing the trends of in the air over 24 hours and the network topology as well as all the other key network characteristics, we generate the same number of flights as the real number of historical flights in our following investigation. Again, the aircraft mobility model developed is capable of generating any number of flights in line with the request of the designers.

A. The Number of Flights in The Air

In this subsection, we first informally characterize the aircraft mobility model developed. The number of flights in the air over a period of 24 hours is shown in Fig. 8 on December 25th, 2018 and June 29th, 2018. The total number of real flights is given here by the sum of the Top-5 Australian

airlines' historical flights on June 29th, 2018 and December 25th, 2018, respectively. In order to validate our mobility model, we artificially generate the same number of flights as the real historical flights on June 29th, 2018 and December 25th, 2018, respectively.

Explicitly, the number of real flights is represented by solid black lines, the flights generated by the Normal distribution are represented by red dotted line, whilst the flights generated by the Weibull distribution are marked by red dashed line. The numbers of flights seen in the figure was generated based on a single realization of a Normal or Weibull process. Observe from Fig. 8(a) and Fig. 8(b) that the sky is quiet before 04:00 on December 25th, 2018 and on June 29th, 2018, respectively. However, there are more and more flights in the air during the daytime. The number of flights reaches its peak between 12:00 and 16:00 for both the real historical flights and for the flights generated by our aircraft mobility model. Furthermore, the number of flights generated by our aircraft mobility model tends to follow the broad trend of the real historical flights. However, there is a drop in the number of real flights at 11:00 both on December 25th, 2018 and on June 29th, 2018.

B. Physical Layer Performance

In order to analyse and validate our aircraft mobility model at the physical layer, we select 12:00 as a specific representative time. The analysis and validation at other times of the day follow the same methodology. Explicitly, let us take a glimpse of the topology and link connection quality at 12:00 both on December 25th, 2018 and on June 29th, 2018, respectively. The link quality is closely tracked by the ACM models shown in Table III. Explicitly, if the distance between a pair of communicating aircraft is higher than 740.8 km, we say there is no available ACM mode; if the distance is longer than 500 km but shorter than or equal to 740.8 km, the ACM mode has a throughput = 0.459 [bps/Hz]; if it is longer than 350 km but shorter than or equal to 500 km, the throughput is 1.000 [bps/Hz]; if it is longer than 200 km but shorter than or equal to 350 km, the throughput is 1.322 [bps/Hz]; if the distance is longer than 110 km but shorter than or equal to 200 km, the throughput is 1.809 [bps/Hz]; if it is longer than 40 km but shorter than or equal to 110 km, the throughput is 2.197 [bps/Hz]; if it is longer than 25 km but shorter than or equal

to 40 km, the throughput is 2.747 [bps/Hz]; if it is longer than 5.56 km but shorter than or equal to 25 km, the throughput is 3.197 [bps/Hz]. Here we use 5.56 km as the minimum safe distance between two aircraft.

As a specific example at 12:00 both on December 25th, 2018 and on June 29th, 2018, the network topology can be readily illustrated by extracting each flight's latitude, longitude and altitude. To provide a glimpse of the location distribution over Australia, we plot the associated 2D topology in Fig. 9 by exploiting the latitude and longitude of each flight. By looking at Fig. 9, we can see that most of the flights cluster near the Eastern coast's airspace both for the real historical flights and for the flights generated by our aircraft mobility model using both the Normal and the Weibull distribution. Note that a slight topology difference can be observed by comparing the real historical flights and the flights generated by our aircraft mobility model, since there is some randomness both in the number of flights generated at a specific time and in their location distribution over the airspace. Nevertheless, the flights generated by both distributions exhibit visual similarity to the real historical flights in terms of topology.

Furthermore, the histograms of the ACM modes at 12:00 on December 25th, 2018 and on June 29th, 2018 are shown in Fig. 10(a) and Fig. 10(b), respectively, which provide a more quantitative comparison between our aircraft mobility model and the real historical flights. We can see from Fig. 10(a) and Fig. 10(b) that most links use ACM mode 1 of Table III, both on December 25th, 2018 and on June 29th, 2018. The number of high-throughput links becomes less and less upon increasing the ACM mode index.

C. Network Layer Performance

In this subsection, we investigate the achievable network layer performance in terms of the ratio of successful connections to a GS, the number of accessible GSs, the minimum number of hops leading to a GS, the area SE defined in bps/Hz/million km^2 [40] and link life-time. In our investigations, there are 15 GSs placed at 15 airports distributed over Queensland, New Southwales, Victoria, South Australia, Western Australia and Northern Territory. By jointly considering their geographical distribution and the busiest of airports, the 15 GSs are placed at Sydney Airport, Melbourne
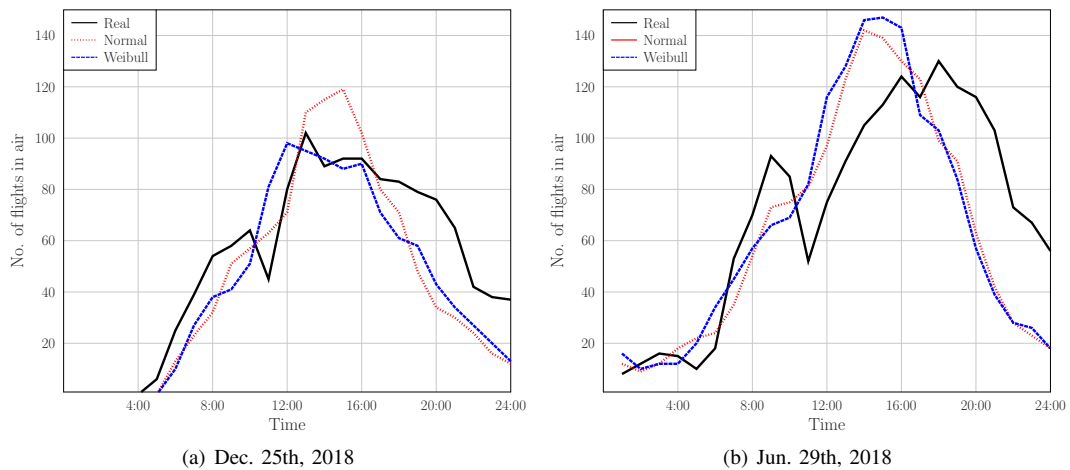

(a) Dec. 25th, 2018   (b) Jun. 29th, 2018

Fig. 8. The number of flights in air over 24 hours in the Australian airspace

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2021.3104118, IEEE Transactions on Vehicular Technology
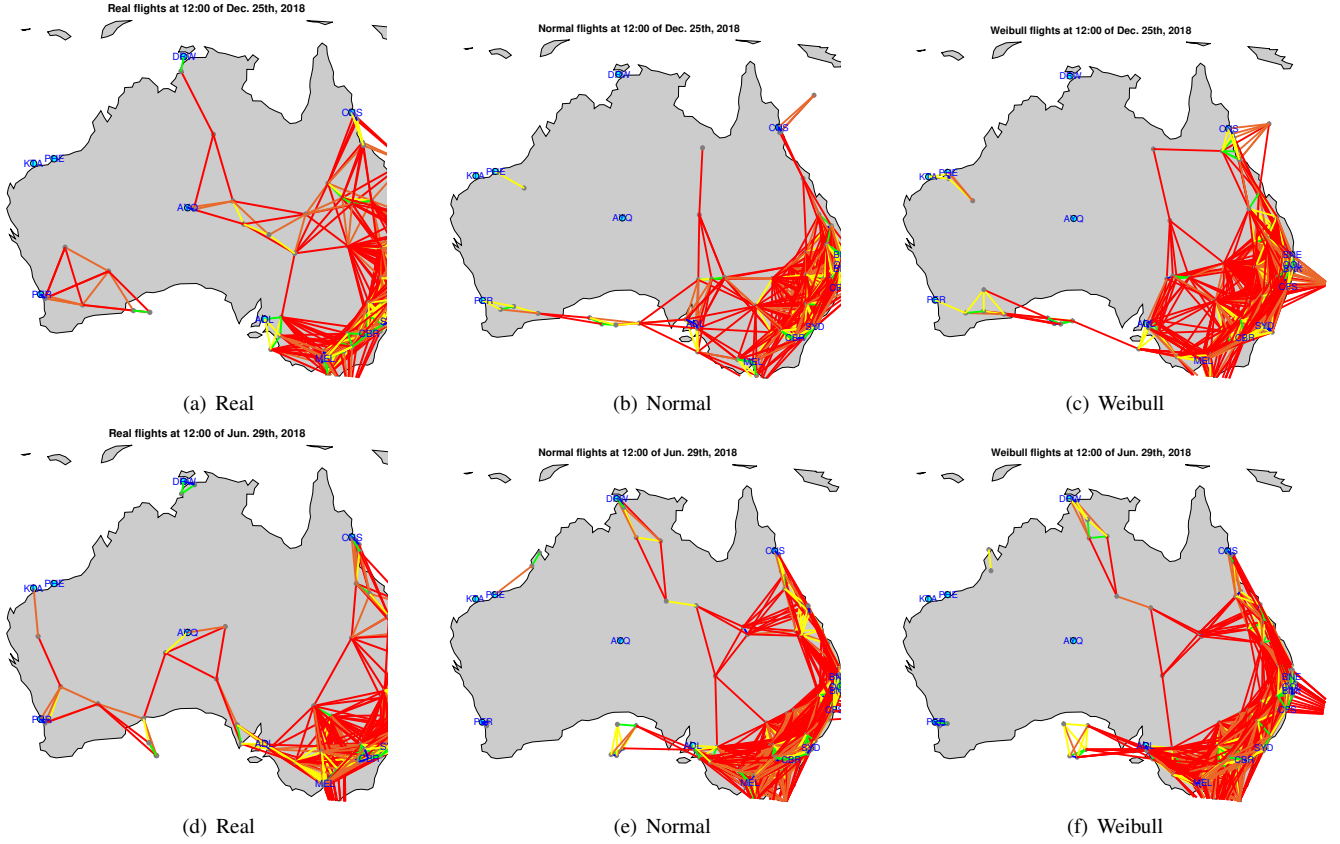
11

Fig. 9. The topology with link connection quality on December 25th, 2018 and on June 29th, 2018, respectively. The subfigures (a)-(c) are topologies with link connection quality on December 25th, 2018. The subfigures (d)-(f) are topologies with link connection quality on December 25th, 2018.
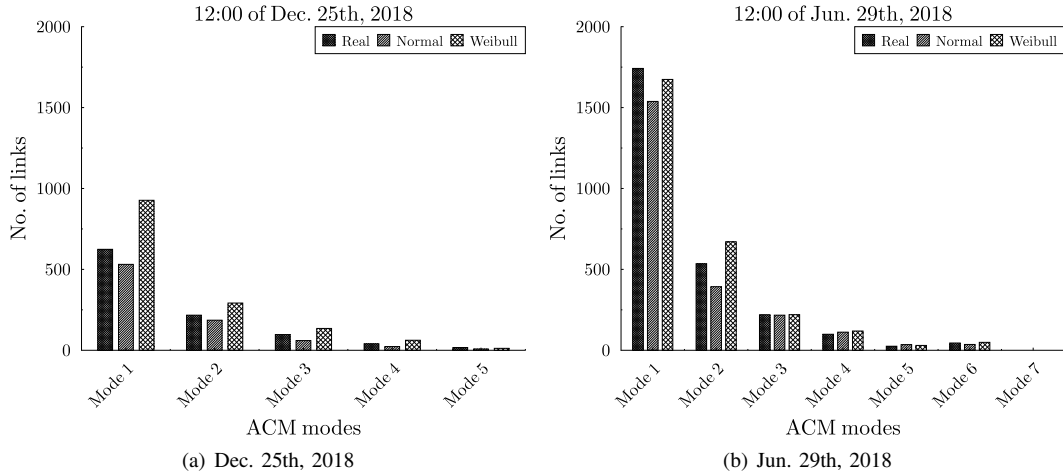

Fig. 10. The histograms of the ACM modes at 12:00 on December 25th, 2018 and on June 29th, 2018, respectively, where the topology of the flights is shown in Fig. 9.

Airport, Brisbane Airport, Perth Airport, Adelaide Airport, Gold Coast Airport, Cairns Airport, Collarenebri Airport, Hobart International Airport, Darwin International Airport, Karratha Airport, Port Hedland International Airport, Ballina Byron Gateway Airport, Coffs Harbour Airport and Ayers Rock Airport (Connellan Airport).

First of all, we investigate whether a specific aircraft in the air is capable of connecting to a GS, which means that it is capable of accessing the Global Internet. The successful

connection ratio R_{con} is defined as

$$R_{con} = \frac{N_{con}}{N_{all}}, \tag{20}$$

where N_{con} is the number of aircraft that is capable of connecting to at least one GS and N_{all} represents the total number of aircraft in the air.

The successful connection ratios over 24 hours on December 25th, 2018 and on June 29th, 2018 are shown in Fig. 11(a) and Fig. 11(b), respectively. Observe from Fig. 11(a) that the successful connection ratios are typically above 90% from
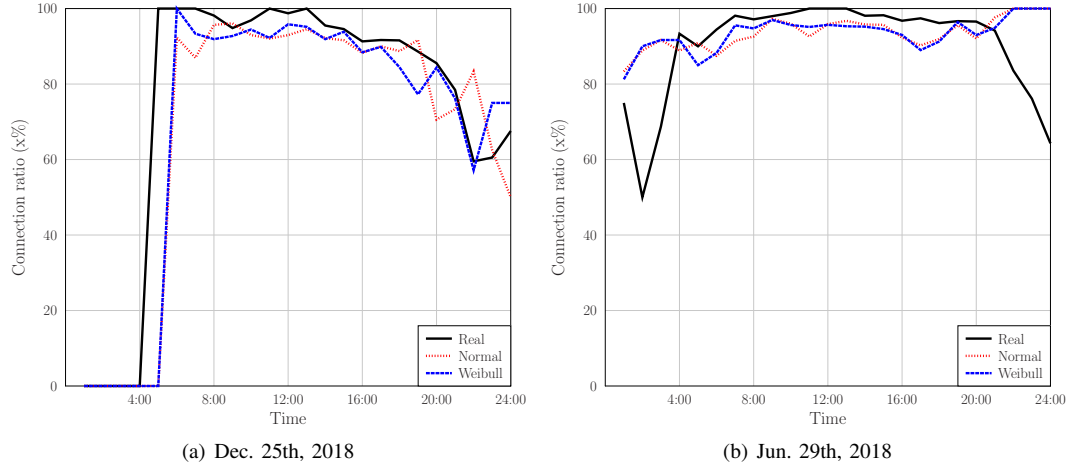
(a) Dec. 25th, 2018
(b) Jun. 29th, 2018

Fig. 11. The ratio of having a connection with any of the ground stations
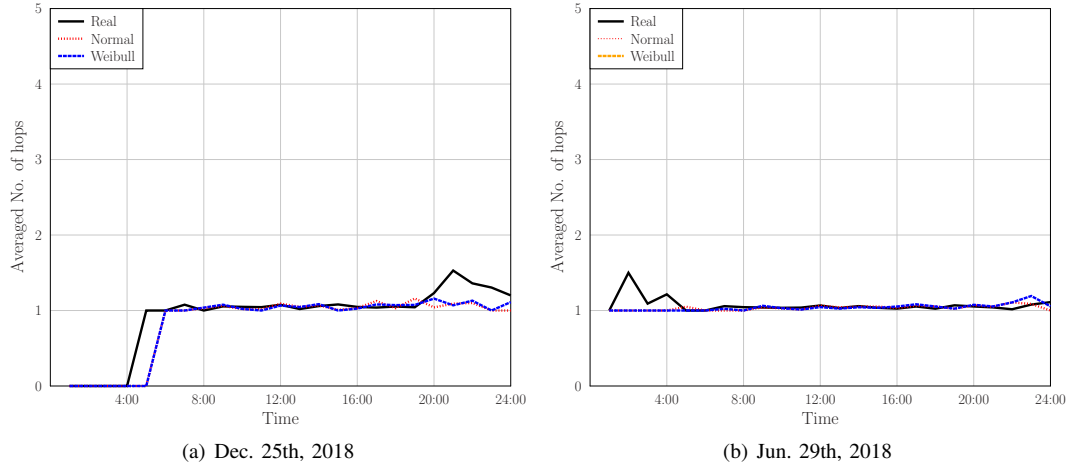

(a) Dec. 25th, 2018
(b) Jun. 29th, 2018

Fig. 12. The average number of hops to a ground station

08:00 to 17:00 on December 25th, 2018, when there are more flights in air during that period of the day. By contrast, the successful connection ratios are typically above 96% from 08:00 to 20:00 on June 29th, 2018, since there are more flights during that period of the day. Furthermore, we conclude that the successful connection ratios have been significantly improved compared to December 25th, 2018, again because there are more flights on June 29th, 2018. We can also observe that the successful connection ratio of the flights generated is capable of closely tracking that of the real historical flights.

Furthermore, the averaged number of hops required by an aircraft for accessing any of the GS is depicted in Fig. 13. Explicitly, Fig. 13(a) depicts the result on December 25th, 2018, while Fig. 13(b) on June 29th, 2018, respectively. Typically, an aircraft is capable of accessing a GS by a single hop both on December 25th, 2018 and on June 29th, 2018, especially during the period between 5:00 and 20:00, which indicates that our placement of the 15 GSs across Australia provides a good coverage for the flights in the Australian airspace.

The achievable maximum area spectral efficiency (ASE) defined by Alouini and Goldsmith [40] is investigated in Fig. 14, whilst the corresponding averaged number of hops

is investigated in Fig. 16. Explicitly, the ASE is defined as the total achievable end-to-end spectral efficiency of all links divided by the total area of Australia. The achievable end-to-end throughput of an aircraft is defined as the spectral efficiency of an aircraft accessing any of the GSs. The area of Australia is given by 7.692 million km^2. By recalling Fig. 8(a) and Fig. 8(b), we can see that the area spectral efficiency is higher, when there are more flights in the air both on December 25th, 2018 and on June 29th, 2018, respectively. The flights generated both by the Normal and by the Weibull distributions are capable of approximating the maximum ASE achieved by the real historical flights.

The average end-to-end throughput was investigated in Fig. 15. Explicitly, Dijkstra's algorithm described in *Section IV* is invoked for finding the routing path of a targeted aircraft with the objective of achieving maximum end-to-end throughput. The end-to-end throughput is limited by the specific link in the routing path that has the minimum link throughput. Furthermore, the distance-based ACM scheme presented *Section III* is also invoked for quantifying the achievable throughput of each link. We can see from Fig. 15(a) that the flights generated by the Normal and Weibull distributions are capable of closely approaching the end-to-end throughput
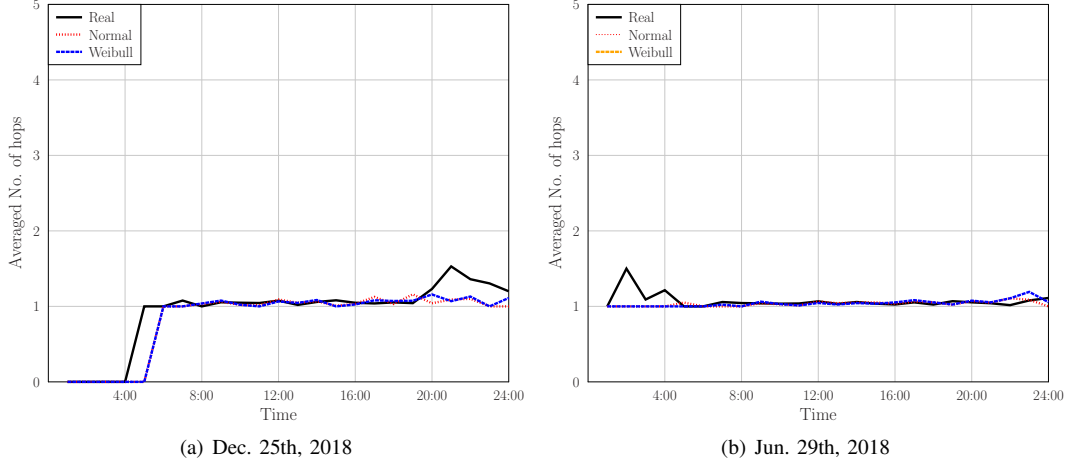
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2021.3104118, IEEE Transactions on Vehicular Technology

13

(a) Dec. 25th, 2018

(b) Jun. 29th, 2018

Fig. 13. The average number of hops to a ground station

(a) Dec. 25th, 2018

(b) Jun. 29th, 2018
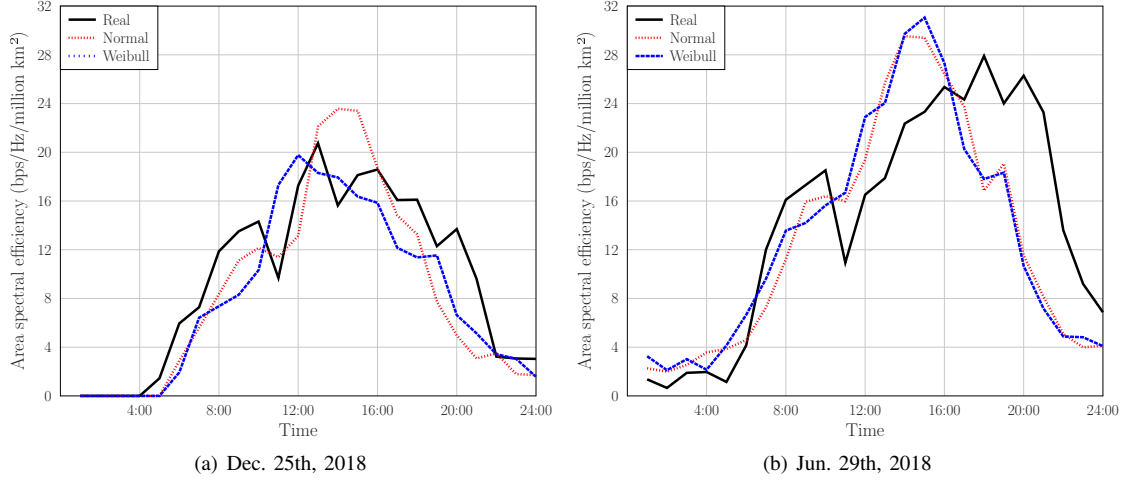
Fig. 14. The area spectral efficiency attained, when having the minimum number of hops to any GS
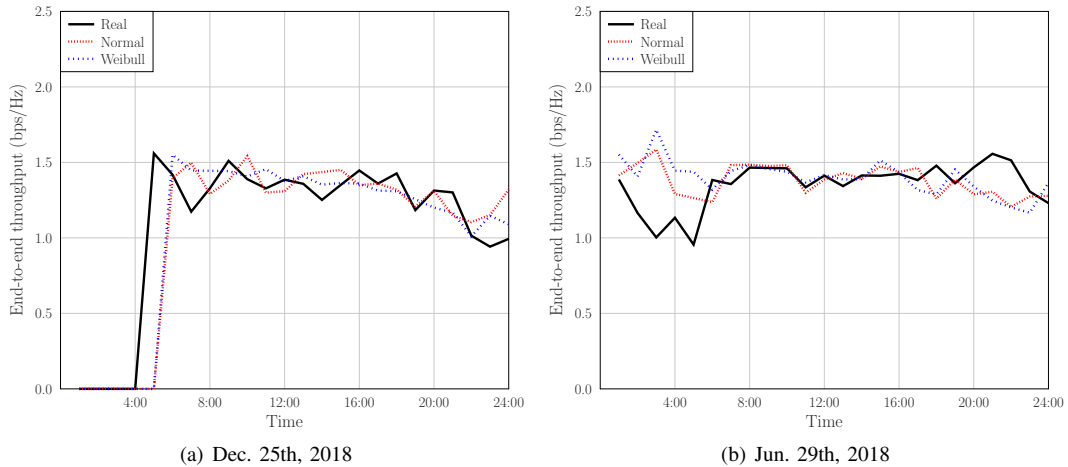

(a) Dec. 25th, 2018

(b) Jun. 29th, 2018

Fig. 15. The average throughput of routing paths aiming for achieving maximum end-to-end throughput.
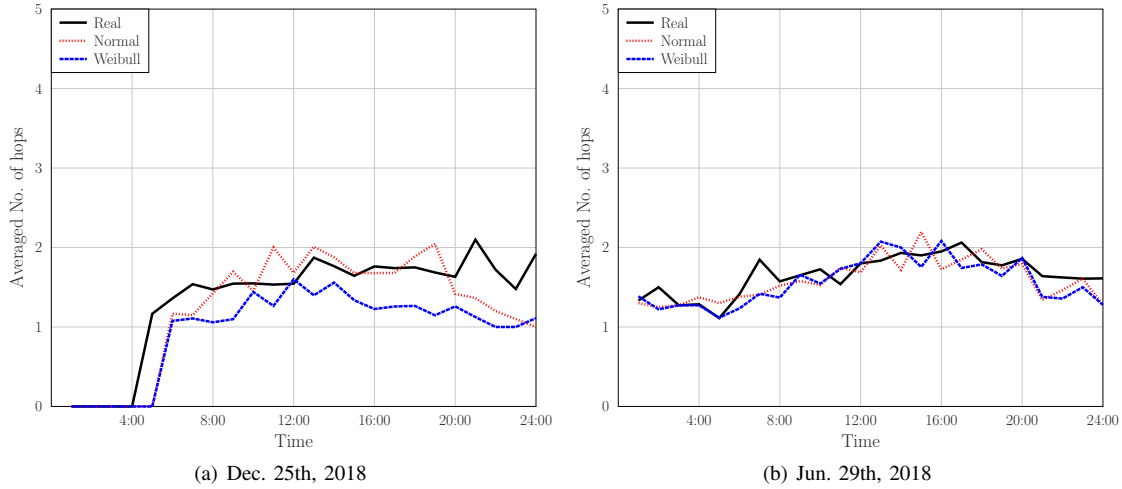
Fig. 16. The average number of hops, when aiming for achieving maximal end-to-end throughput

achieved by the real historical flights between 8:00 and 21:00 on December 25th, 2018, when there are more flights in air by recalling Fig. 8(a). Similarly, the flights generated by the Normal and Weibull distributions are also capable of approaching the end-to-end throughput attained by the real flights from 5:00 to 17:00 on June 29th, 2018 by observing Fig. 15(b).

The corresponding average number of hops associated with achieving maximal end-to-end throughput on December 25th, 2018 and June 29th, 2018 are portrayed in Fig. 16(a) Fig. 16(b), respectively. We can see that an aircraft may need more hops for achieving a higher end-to-end throughput than in Fig. 13, which depicts the averaged number of hops required for accessing any of the GSs. Furthermore, by observing Fig. 16(a) Fig. 16(b), we can also see that the flights generated by the semi-stochastic aircraft mobility model is capable of tracking the trend of the number of hops imposed over 24 hours.

The link life-time is defined as the duration of time for a link to be maintained between a pair of aircraft, which is one of the most important metrics in ad-hoc networks. These results recorded for the real historical flights and for the flights generated by our aircraft mobility model are shown in Fig. 17. Explicitly, Fig. 17 portrays the histogram of link life-time at intervals of $1/2$ hour, whilst Fig. 17 portrays their complementary cumulative distribution function (CCDF) statistics. Observe from Fig. 17 that most of the links have a link life-time of $3/4$ hour and/or $1\frac{1}{4}$ hours both for the real historical flights and the flights generated by the Normal and Weibull distributions. There are almost no flights having a 2-hour link life-time.

Furthermore, the link life-time may be improved by conceiving sophisticated routing protocols. Intuitively, a pair of aircraft having the same flying direction and similar speed will have a long-lasting link connectivity, where the estimated level-crossing rate of the link can be invoked as a metric for finding long-lasting routing paths.

VI. DISCUSSIONS ABOUT THE APPLICABILITY TO OTHER SCENARIOS

The methodology presented in *Section II* for our developing semi-stochastic aircraft mobility model is exemplified by the Australian scenario and its performance is validated by Australian flight data. However, it is also applicable to other scenarios, such as Europe, the North-Atlantic (NA), the United States (US) and China, provided that their historical flight data is available. Having said that, naturally, the best fitting distribution used for artificially generating new flights is scenario-dependent. To elaborate, there are three typical airspaces, namely sparse airspaces over populated areas, dense airspaces over populated areas and trans-ocean airspaces. The sparse airspaces over populated areas may be represented by the Australian airspace, the dense airspaces over populated areas are typical for Europe, US and China, whilst the trans-ocean airspace may be represented by the NA airspace or Trans-Pacific airspace. In the following, we will briefly highlight the application of our semi-stochastic aircraft mobility model both to the European scenario and to the NA scenario, which represent the dense airspace over populated areas and the trans-ocean airspace, respectively.

Explicitly, the Top-5 airlines transatlantic flights were exploited for our investigation of the NA scenario, which are Delta Airlines, United Airlines, American Airlines, British Airways and Lufthansa. Because there is significant time difference between flights taking off from Europe and the US, we investigate their statistical characteristics separately. Since there are too many airlines in Europe, if we only consider top-5 airlines' flights, they cannot cover all flight route paths across Europe. Hence, we choose all of the departure and arrival flights of the top-5 busiest airports for European scenario. The top-5 busiest airports selected are Heathrow Airport, UK; Charles de Gaulle Airport, France; Amsterdam's Schiphol Airport, Netherlands; Istanbul's Ataturk Airport, Turkey; Frankfurt Airport, Germany. The goodness-of-fit values of the six representative hypothesis distributions investigated are summarized in Table V.

Again, without considering the non-parametric Kernel density estimation, the Gamma distribution has the best the best
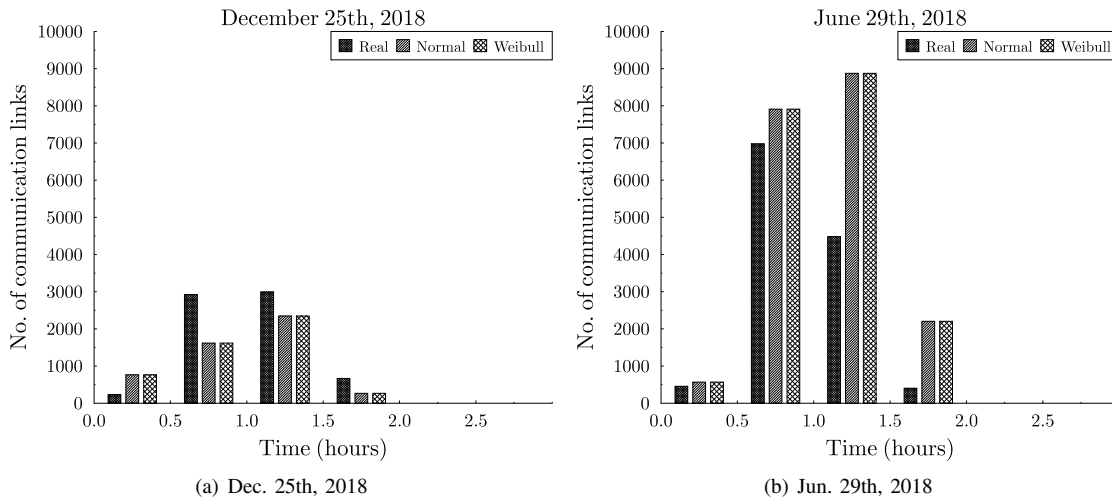
(a) Dec. 25th, 2018

(b) Jun. 29th, 2018

Fig. 17. The histogram of link life time in Australia

goodness-of-fit for the NA-EU side both on June 29th 2018 and on December 25th 2018, whilst the Student-t distribution has the best goodness-of-fit for the NA-US side both on June 29th 2018 and on December 25th 2018. By contrast, the Normal distribution has the best goodness-of-fit for Europe scenario both on June 29th 2018 and on December 25th 2018. Furthermore, the best fitting parameters of the Gamma distribution, Student-t distribution and Normal distribution are summarised in Table VI, which can be used for generating near-realistic flight data for both the NA scenario and European scenario by recalling the semi-stochastic aircraft mobility models, respectively.

VII. CONCLUSIONS

Since there is a paucity of literature on aircraft mobility models for generating near-realistic flight data for investigating aeronautical networks, we developed a semi-stochastic aircraft mobility model based on large-scale real historical flights acquired on the quietest day and busiest day of 2018. The proposed aircraft mobility model is capable of generating near-realistic mobility that captures the statistical features of the number of aircraft in air, their motion trajectory, topology, link quality as well as network layer performance. The achievable link quality of the physical layer was investigated by relying on our distance-based ACM that is specifically designed for aeronautical communications. In order to investigate the network layer performance, we developed a single-source-to-multiple-destination routing optimization scheme based on Dijkstras algorithm. Furthermore, we analysed and validated our aircraft mobility model by investigating and comparing the statistical link quality, end-to-end -throughput, link connection ratio to any ground station, the number of hops as well as the link life-time. Finally, the key statistical parameters of the distribution used for generating artificial flight data for dense airspaces represented by European airspace and for trans-ocean airspaces represented by the NA airspace were characterized. Our future research will consider the multi-component Pareto-optimization of similar networks.

REFERENCES

[1] J. Zhang, T. Chen, S. Zhong, J. Wang, W. Zhang, X. Zuo, R. G. Maunder, and L. Hanzo, "Aeronautical ad-hoc networking for the Internet-above-the-clouds," *Proceedings of the IEEE*, vol. 107, no. 5, pp. 868–911, 2019.
[2] X. Huang, J. A. Zhang, R. P. Liu, Y. J. Guo, and L. Hanzo, "Airplane-aided integrated networking for 6G wireless: Will it work?" *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 84–91, September 2019.
[3] S. Batabyal and P. Bhaumik, "Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1679–1707, 2015.
[4] X. Wu, H. R. Sadjadpour *et al.*, "Link dynamics in manets restricted node mobility: Modeling and applications," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4508–4517, September 2009.
[5] G. H. Mohimani, F. Ashtiani, A. Javanmard, and M. Hamdi, "Mobility modeling, spatial traffic distribution, and probability of connectivity for sparse and dense vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1998–2007, May 2009.
[6] H. Wang, R. P. Liu, W. Ni, W. Chen, and I. B. Collings, "VANET modeling and clustering design under practical traffic, channel and mobility conditions," *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 870–881, March 2015.
[7] H. Peng, A. Razi, F. Afghah, and J. Ashdown, "A unified framework for joint mobility prediction and object profiling of drones in UAV networks," *Journal of Communications and Networks*, vol. 20, no. 5, pp. 434–442, October 2018.
[8] M. S. Murty and M. V. Das, "Performance evaluation of manet routing protocols using reference point group mobility and random waypoint models," *International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC)*, vol. 2, no. 1, pp. 33–43, March 2011.
[9] F. K. Karnadi, Z. H. Mo, and K.-c. Lan, "Rapid generation of realistic mobility models for VANET," in *2007 IEEE wireless communications and networking conference*. Hong Kong, China: IEEE, March 2007, pp. 2506–2511.
[10] K. Kumari, S. Maakar, and B. Sah, "A brief survey of mobility model for FANET," in *Proceedings of national conference on innovative trends in computer science engineering*, Brcmcet, Bahal, April 2015, pp. 106–108.
[11] N. Lin, F. Gao, L. Zhao, A. Al-Dubai, and Z. Tan, "A 3D smooth random walk mobility model for FANETs," in *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*. Zhangjiajie, China: IEEE, August 2019, pp. 460–467.
[12] D. Medina and F. Hoffmann, *The airborne internet*. InTech, September 2011.
[13] A. Tiwari, A. Ganguli, A. Sampath, D. S. Anderson, B.-H. Shen, N. Krishnamurthi, J. Yadegar, M. Gerla, and D. Krzysiak, "Mobility aware routing for the airborne network backbone," in *IEEE Military Communications Conference*, San Diego, USA, November 2008, pp. 1–7.
[14] D. He, W. Sun, and L. Shi, "The novel mobility models based on spiral line for aerial backbone networks," *IEEE Access*, vol. 8, pp. 11 297–11 314, January 2020.

TABLE V
GOODNESS OF FIT OF STATISTICAL MODEL FOR THE TAKEOFF TIME FOR NA SCENARIO AND EU SCENARIO

Representative dates	Scenarios	Weibull	**Gamma**	Kernel	Normal	Poisson	Student-t
June 29th, 2018	NA-EU side	0.1380	**0.0972**	0.0982	0.1061	0.1208	0.1056
	NA-US side	0.1467	0.2173	0.0926	0.1628	0.2007	**0.0857**
	Europe	0.0760	0.0894	0.0313	**0.0671**	0.1595	0.0672
December 25th, 2018	NA-EU side	0.1717	**0.1154**	0.1237	0.1310	0.1269	0.1428
	NA-US side	0.1478	0.2148	0.1245	0.1741	0.2083	**0.1177**
	Europe	0.0776	0.1064	0.0468	**0.0755**	0.1574	0.0755

TABLE VI
THE BEST FITTING DISTRIBUTION PARAMETERS FOR NORTH-ATLANTIC SCENARIO AND EUROPE SCENARIO

Representative dates	Scenarios	Distribution	Parameters	Values
June 29th, 2018	NA-EU side	Gamma	Shape parameter a_g	14.51
			Scale parameter b_g	1.12
	NA-US side	Student-t	Location parameter μ_t	16.49
			Scale parameter σ_t	2.06
			Shape parameter ν_t	2.76
	Europe	Normal	Mean μ_J	12.85
			Variance σ_J	5.17
December 25th, 2018	NA-EU side	Gamma	Shape parameter a_g	11.42
			Scale parameter b_g	1.02
	NA-US side	Student-t	Location parameter μ_t	17.93
			Scale parameter σ_t	2.02
			Shape parameter ν_t	3.00
	Europe	Normal	Mean μ_J	13.29
			Variance σ_J	5.19

[15] J. P. Rohrer, E. K. Cetinkaya, H. Narra, D. Broyles, K. Peters, and J. P. Sterbenz, "AeroRP performance in highly-dynamic airborne networks using 3D Gauss-Markov mobility model," in *2011 Military Communications Conference*. Baltimore, USA: IEEE, November 2011, pp. 834–841.

[16] W. Wang, X. Guan, B. Wang, and Y. Wang, "A novel mobility model based on semi-random circular movement in mobile ad hoc networks," *Information Sciences*, vol. 180, no. 3, pp. 399–413, 2010.

[17] P. K. Sharma and D. I. Kim, "Random 3D mobile UAV networks: Mobility modeling and coverage probability," *IEEE Transactions on Wireless Communications*, vol. 18, no. 5, pp. 2527–2538, May 2019.

[18] D. Broyles and A. Jabbar, "Design and analysis of a 3-D Gauss-Markov model for highly dynamic airborne networks." San Diego, USA: International Foundation for Telemetering, November 2010, pp. 1–7.

[19] J. Xie, Y. Wan, J. H. Kim, S. Fu, and K. Namuduri, "A survey and analysis of mobility models for airborne networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1221–1238, 2014.

[20] J. Li, L. Lei, W. Liu, Y. Shen, and G. Zhu, "An improved semi-Markov smooth mobility model for aeronautical ad hoc networks," in *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing*. Limassol, Cyprus: IEEE, August 2012, pp. 1–4.

[21] Y. Wan, K. Namuduri, Y. Zhou, and S. Fu, "A smooth-turn mobility model for airborne networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3359–3370, 2013.

[22] J. Xie, Y. Wan, B. Wang, S. Fu, K. Lu, and J. H. Kim, "A comprehensive 3-dimensional random mobility modeling framework for airborne networks," *IEEE Access*, vol. 6, pp. 22 849–22 862, 2018.

[23] T. Gräupl, "FACTS2: Extended simulation framework for ATM communication demand analysis of Europe," in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*. St. Petersburg, USA: IEEE, November 2017, pp. 1–8.

[24] ——, "Validating the FACTS2 air traffic simulation framework," in *2017 Integrated Communications, Navigation and Surveillance Conference (ICNS)*. Herndon, USA: IEEE, April 2017, pp. 1D4–1.

[25] D. Salomon, *Curves and surfaces for computer graphics*. Springer Science & Business Media, 2007.

[26] Z. Drezner, O. Turel, and D. Zerom, "A modified Kolmogorov–Smirnov test for normality," *Communications in Statistics–Simulation and Computation®*, vol. 39, no. 4, pp. 693–704, 2010.

[27] E. Haas, "Aeronautical channel modeling," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 2, pp. 254–264, March 2002.

[28] Z. Ma, B. Ai, R. He, G. Wang, Y. Niu, and Z. Zhong, "A wideband non-stationary air-to-air channel model for UAV communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1214–1226, February 2019.

[29] J. Zhang, S. Chen, R. G. Maunder, R. Zhang, and L. Hanzo, "Adaptive coding and modulation for large-scale antenna array-based aeronautical communications in the presence of co-channel interference," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1343–1357, 2017.

[30] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell system technical journal*, vol. 47, no. 6, pp. 957–1000, 1968.

[31] P. Hoeher, "A statistical discrete-time model for the WSSUS multipath channel," *IEEE Transactions on vehicular technology*, vol. 41, no. 4, pp. 461–468, 1992.

[32] J. Zhang, S. Chen, R. G. Maunder, R. Zhang, and L. Hanzo, "Regularized zero-forcing precoding-aided adaptive coding and modulation for large-scale antenna array-based air-to-air communications," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 2087–2103, 2018.

[33] D. Wang, Y. Wang, S. Dong, G. Huang, J. Liu, and W. Gao, "On delay-aware resource control with statistical QoS provisioning by dual connectivity in heterogeneous aeronautical network," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2915–2927, March 2020.

[34] O. S. Oubbati, A. Lakas, F. Zhou, M. Güneş, and M. B. Yagoubi, "A survey on position-based routing protocols for Flying Ad hoc Networks (FANETs)," *Vehicular Communications*, vol. 10, pp. 29–56, October 2017.

[35] B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV assistance paradigm: State-of-the-art in applications and challenges," *Journal of Network and Computer Applications*, vol. 166, pp. 1–44, September 2020.

[36] O. S. Oubbati, M. Atiquzzaman, P. Lorenz, M. H. Tareque, and M. S. Hossain, "Routing in flying ad hoc networks: survey, constraints, and future challenge perspectives," *IEEE Access*, vol. 7, pp. 81 057–81 105, June 2019.

[37] D. Medina, F. Hoffmann, F. Rossetto, and C.-H. Rokitansky, "A geographic routing strategy for north atlantic in-flight internet access via airborne mesh networking," *IEEE/ACM Transactions on Networking*, vol. 20, no. 4, pp. 1231–1244, August 2011.

[38] Q. Luo and J. Wang, "Multiple QoS parameters-based routing for civil aeronautical ad hoc networks," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 804–814, June 2017.

[39] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2009.

[40] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Transactions on vehicular technology*, vol. 48, no. 4, pp. 1047–1066, July 1999.