Distribution of Longitudinal Speed Prediction Error of ADS-C System

Masato Fujita
Air Traffic Management Department
Electronic Navigation Research Institute
Chofu, JAPAN
m-fujita@enri.go.jp

Abstract— **The number of aircraft flying in oceanic airspaces is growing. To accommodate the traffic growth, the reduction of separation minimum for Automatic Dependent Surveillance − Contract (ADS-C) aircraft is required. However, the reduction of the separation minimum increases the collision risk of aircraft and the safety assessment prior to the reduction is expected. The probability distribution model of the longitudinal speed prediction error is a key parameter of the collision risk formula for the longitudinal separation minimum under ADS-C. In this paper, the empirical distribution of the longitudinal speed prediction error of aircraft in North Pacific routes is provided. Using Peak over Threshold (POT) technique, we found the distribution model which is appropriate for the risk estimation.**

Keywords-component; Automatic Dependent Surveillance – Contract (ADS-C), Longitudinal Speed Prediction Error, Peak Over Threshold, Collision Risk

I. INTRODUCTION

NOPAC (NOrth PACific) route system (Fig.1) is the most congested oceanic ATS route system in Fukuoka FIR. The number of aircraft flying NOPAC route system is growing. To accommodate the traffic growth, the reduction of separation minima is expected. The 50NM longitudinal separation minimum for ADS-C (Automatic Dependent Surveillance - Contract) aircraft has been implemented sequentially beginning from R220 and R580. In near future, the 30NM longitudinal separation minimum will be implemented.

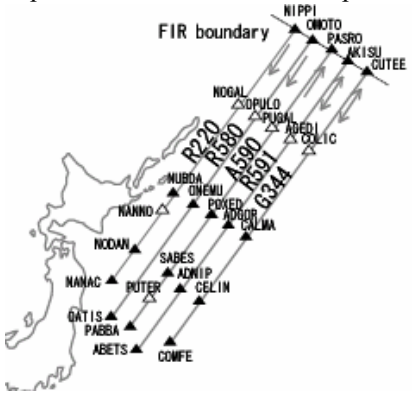

Figure 1. NOPAC route system

An aircraft under ADS-C circumstance transmits their position periodically. In Fukuoka FIR, the reporting interval is 1600 seconds in usual and 320 seconds in the case of strategic lateral offset. Under radar surveillance, position information is available in the order of seconds. Since the position information is rarely available under ADS-C circumstance, the prediction of trajectory is indispensable for surveillance.

For the quantitative estimation of mid-air collision risk of aircraft, the expected number of fatal accidents is often utilized as a risk indicator. It is called a collision risk. When the estimated collision risk does not exceed the target level of safety which is determined prior to the safety assessment, the situation is considered to be safe. When the separation minima are reduced, the collision risk increases. Hence, the safety assessment is required to confirm that the airspaces remain safe even under the reduced separation minima. (See [1].)

An aircraft pair collides if and only if they overlap in longitudinal, lateral and vertical dimension. Hence, the longitudinal overlap probability which is the probability that a pair of aircraft overlap in the longitudinal dimension should be estimated in the safety assessment of reduced longitudinal separation minimum. It is calculated using the probability distribution of aircraft position error due to aircraft navigation capability and the probability distribution of speed prediction error which causes from the position prediction performance of on-board systems and the interpolation performance of ground systems. (See [2] and [3].)

This paper gives the empirical distribution of longitudinal speed prediction error and the curve fitted to the empirical distribution applying POT (Peak over Threshold) technique in Extreme Value Theory. (See [4].)

II. CONCEPT OF ADS -C

Under ADS-C circumstance, ground stations transmit a message which tells the required type of downlink messages and the frequency of downlinks. It is called a contract message. An ADS-C aircraft downlink the required messages automatically as it is indicated in the contract messages. Downlink is executed periodically (periodic report), when the

ISBN: 978-0-615-20720-9

event (lateral deviation, vertical rate change, waypoint change, altitude range change) occurs (event report), or a one-time-only report required by ATC (demand report). Downlink ADS messages are classified into basic messages and the other optional messages. A basic message contains stamped time, current position etc. Optional messages, for instance, provide the ground speed and direction at the reporting time, the location of the next waypoint and its estimated time of arrival and the predicted position at some future time instance. A ground ATC system, which is called ODP (Oceanic Data Processing system) in Japan, interpolates (and extrapolates) to predict the aircraft position from the optional messages till it receives next report.

In Japanese system, predicted route group messages, intermediate projected intended group messages and fixed projected intent group messages are utilized for the prediction of aircraft position at the ground system. The first message provides the location of the next waypoint over which the aircraft is passing and the estimated time of arrival. The third gives the predicted position at some future instance. Japanese system requires ADS aircraft to send the predicted position 37 minutes later. When an aircraft intends to change its speed or direction within 37 minutes, intermediate project intended group messages are coupled to inform when and where the speed and direction are changed.

III. DERIVATION OF EMPRICAL MODEL

A. Identification of ADS-C messages of aircraft on NOPAC

An ADS message includes a position report and the predicted position of aircraft. However, it does not contain the route name on which the aircraft intended to fly. To identify which route each aircraft flies on, FDPS (Flight Data Processing System) data set was utilized.

All flights in Fukuoka FIR are saved in FDPS data with their call signs, aircraft types, the departure and destination airports, the original flight plans, the waypoints over which the aircraft flew, the time instance when aircraft flew over the waypoints etc. All flights of NOPAC routes are identified by FDPS data set.

The ADS-C and ATS Facilities Notification (AFN) data set during September 1st 2005 to August 31st 2006 in the format of [5], [6] and [7] were provided by Kobe Aeronautical Satellite Center (the data on November 18th 2005 and from January 17th 2006 to February 9th 2006 could not be collected). The FDPS data set in the same period was provided by ATM Center.

The AFN procedure enables an ATS facility to become aware of an aircraft's data link capability and provides an exchange of address information. AFN messages are transmitted when an aircraft enters into a region where a data link service is provided by a service provider and when an aircraft is placed under the control of an adjacent ATS facility. All ADS-C messages transmitted in one flight are wedged by AFN messages and ADS-C disconnect messages in chronological order. Since the AFN message contains the aircraft registration number, for every ADS-C message, the registration number of aircraft which transmits the ADS-C message is identified.

In many cases, the registration number of aircraft which was utilized in a flight is included in FDPS data. The corresponding ADS-C messages were identified using the registration number as the search key. However, in the case where no registration number of an ADS aircraft flying NOPAC is saved in FDPS data, we found the corresponding ADS messages manually with the help of self-developed GUI. (Fig. 2)

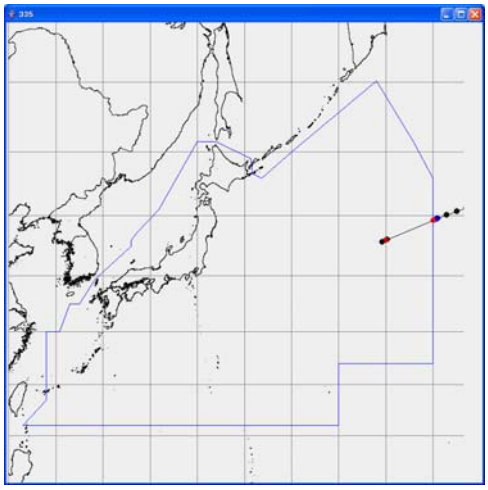

Figure 2. Display of GUI
(Black dots means the periodic reports and
colored dots means the event-driven position reports)

B. Definition of Longitudinal Speed Prediction Errors

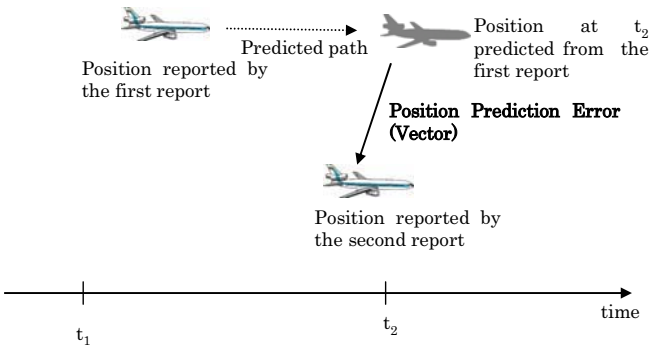

Figure 3. Definition of Speed Prediction Error

Consider successive two ADS-C messages transmitted by a single aircraft. Let t_1 be the time instance when the first message was transmitted and let t_2 be the time instance when the second message was transmitted. The position prediction error is defined as the difference of the reported position at t_2

ISBN: 978-0-615-20720-9

from the predicted position at t_2 estimated from the ADS-C message transmitted at t_1. The speed prediction error is the position prediction error divided by (t_2-t_1). The longitudinal speed prediction error is the projection of the speed prediction error onto the route direction. (Fig. 3)

C. Definition of Longitudinal Speed Prediction Errors

1) Assumption used for the prediction

We briefly state the assumption which is utilized in the calculation of the predicted position.

In the observation period, the basic messages, the predicted route group messages and the fixed projected intent group messages were utilized for the position prediction in ODP. ODP assumes that an aircraft is headed to the point indicated in the predicted route group messages or the fixed projected intent group messages, whose estimated time of arrival is earlier, and in the next step, the aircraft headed to the other point indicated in the messages. We assume the same assumption in this paper.

The earth is assumed to be a sphere and an aircraft is assumed to fly on the great circle in this paper. This is also assumed in ODP.

2) Calculation of the estimated position

We introduce the mathematical algorithm to calculate the predicted position T at M hours later given the reported position P and the predicted position Q at N hours later. Remark that the mathematical algorithm described below is not identical to the one used in ODP.

Let O be the center of the sphere. The aircraft flies with a constant speed on the cross section of the sphere by the plane OPQ. An aircraft which flies $d_S(P,Q)$ by distance in N hours flies M $d_S(P,Q)/N$ by distance in M hours. Here $d_S(P,Q)$ denotes the distance between points P and Q on the sphere given by

$$d_S(P,Q) = 2R\sin^{-1}(d_{3-\dim}(P,Q)/2R) \tag{1}$$

The notation R is the radius of the sphere and $d_{3\text{-}dim}(P,Q)$ is the Euclid distance in the 3-dimensional space.

Let φ be the angle made by half lines OP and OT. By the definition of T, we get the following equation.

$$R\varphi = d_S(P,T) = \frac{M}{N} d_S(P,Q) \tag{2}$$

Let H be the foot of perpendicular of the point O on the line PQ. Let S be the crossing point of the line PQ with the line OT. (Fig. 4) Let θ be the angle made by half lines OP and OQ, then it is given by

$$\theta = 2\sin^{-1}(d_{3-\dim}(P,Q)/2R) \tag{3}$$

Even if the pint is on the segment PH, HQ or right of Q in Fig. 4, the following equation holds true.

$$d_{3\text{-}dim}(P,S) = R\sin(\theta/2) + R\cos(\theta/2)\tan(\varphi-\theta/2) \tag{4}$$

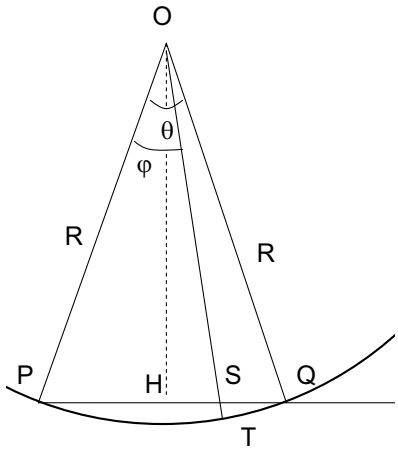

Figure 4. Geometrical Description I

Let $P(x_P, y_P, z_P)$, $Q(x_Q, y_Q, z_Q)$, $S(x_S, y_S, z_S)$ and $T(x_T, y_T, z_T)$ be the coordinates. Then, the following equations are obviously satisfied.

$$\begin{aligned}&(x_S, y_S, z_S) = \\ &(x_P + \frac{d_{3-\dim}(P,S)}{d_{3-\dim}(P,Q)}(x_Q - x_P), y_P + \frac{d_{3-\dim}(P,S)}{d_{3-\dim}(P,Q)}(y_Q - y_P), \\ &\quad z_P + \frac{d_{3-\dim}(P,S)}{d_{3-\dim}(P,Q)}(z_Q - z_P))\end{aligned} \tag{5}$$

$$\begin{aligned}&(x_T, y_T, z_T) \\ &= (\frac{Rx_S}{\sqrt{x_S^2 + y_S^2 + z_S^2}}, \frac{Ry_S}{\sqrt{x_S^2 + y_S^2 + z_S^2}}, \frac{Rz_S}{\sqrt{x_S^2 + y_S^2 + z_S^2}})\end{aligned} \tag{6}$$

By (2) - (6), we get the coordinate of the point T.

3) Calculation of the longitudinal position errors

Let A be the reported position at some time instance and let B be the predicted position at the same time estimated in the methodology described in the previous subsection. Using the coordinates of waypoints, we find the equation of the plane W containing the center O and the route. The author used Gaussian Elimination to find the equation. Let

$$ax+by+cz=0 \tag{7}$$

be the equation.

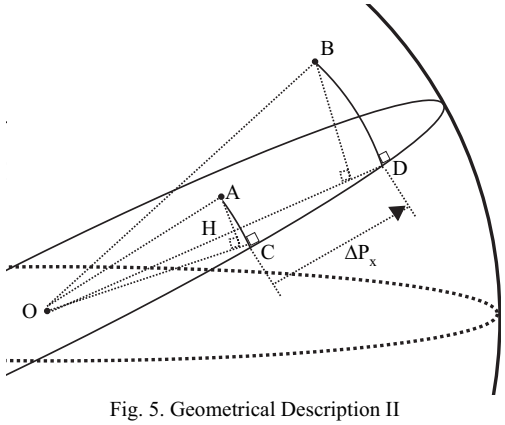

Fig. 5. Geometrical Description II

ISBN: 978-0-615-20720-9

We introduce how to find the coordinate of 'the foot C of perpendicular from A on the route.' Let H be the foot of perpendicular on the plane W. The coordinate of the point H is given by

$$(x_H, y_H, z_H) = \left(x_A - a\frac{ax_A + by_A + cz_A}{a^2+b^2+c^2}, y_A - b\frac{ax_A + by_A + cz_A}{a^2+b^2+c^2}, z_A - c\frac{ax_A + by_A + cz_A}{a^2+b^2+c^2}\right) \quad (8)$$

Hence the coordinate of the point C is given by

$$(x_C, y_C, z_C) = \left(\frac{Rx_H}{\sqrt{x_H^2+y_H^2+z_H^2}}, \frac{Ry_H}{\sqrt{x_H^2+y_H^2+z_H^2}}, \frac{Rz_H}{\sqrt{x_H^2+y_H^2+z_H^2}}\right) \quad (9)$$

We can find the coordinate of 'the foot D of perpendicular from B on the route' in the same way. We define the longitudinal speed prediction error by

$$\Delta P_x = d_S(C,D) \quad (10)$$

if the point C and D line up in the traveling direction of the aircraft, otherwise, it is defined by

$$\Delta P_x = -d_S(C,D) \quad (11)$$

D. Results

We study the longitudinal speed prediction error in the case where aircraft fly 'straight and at a constant speed.' When the event report is transmitted, the aircraft assumes to change its speed, heading or its vertical speed. Hence we only consider the successive ADS reports such that both of them are assumed to be periodic reports. (The basic group report following the contract message is assumed to be a demand report, when the contract message for a demand message is transmitted.)

Fig. 6 shows the time interval of successive periodic reports of aircraft flying in the NOPAC route system. Since the reporting time interval indicated in the contract message is 320 sec and 1600 sec, there are peaks at 6 min and 27 min.

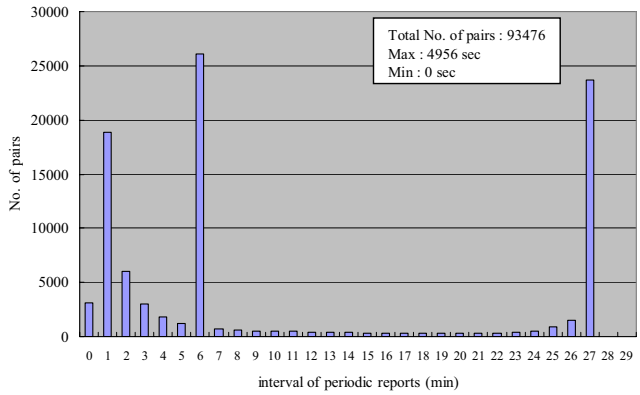

Figure 6. Distribution of intervals of periodic reports

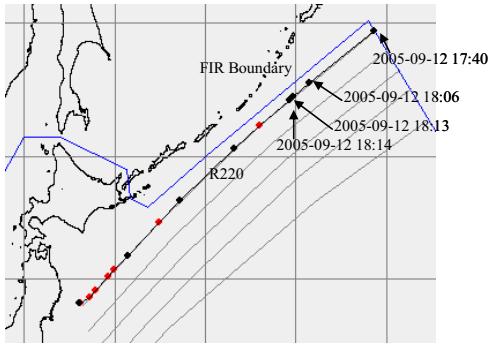

Figure 7. Trajectory of an aircraft with large longitudinal speed prediction error

In the rest of this paper, we only consider the periodic report pairs whose time interval is greater than 26 min and less than or equal to 27 min (right peak of Fig. 6). A few ADS-C reports in this data set are not coupled with fixed intent group. In the case where the estimated time of arrival at the next waypoint described in the predicted route group message is close to the stamped time of the basic message, the longitudinal prediction error is sometimes large in magnitude.

Fig. 7 shows the trajectory of an aircraft flying on R220 in the NOPAC route system. The dots show the position reported via ADS-C. The longitudinal speed prediction error of reports which were transmitted at 17:40 and 18:06 was -668 (knots). It turned out that the basic group transmitted at 17:40 is not coupled with 'fixed projected intent group' and estimated time of arrival given in 'predicted route group' is 10 sec later from the stamped time. The 'predicted route group' reports that the next waypoint is in the east of the reported position in spite of the westbound aircraft. It seems that the aircraft flies by the waypoint; however, the on-board system does not update the next waypoint. Hence the ground system possibly considers that the aircraft is flying in the opposite direction. If a fixed projected intend group is coupled in this case, a system might misunderstand the aircraft heading in a short period. However, the system makes an appropriate prediction based on the fixed projected intent group after a few seconds.

0.034% of basic reports are not coupled with 'fixed projected intent group' and 0.017% of basic reports are coupled with neither 'fixed projected intent group' nor 'predicted route group'.

Fig. 8 shows the empirical distribution of the longitudinal speed prediction errors of periodic report pairs which are coupled with both fixed projected intent group and predicted route group and whose reporting time interval is greater than 26 min and less than or equal to 27 min. There are no incredibly large longitudinal speed prediction errors any more.

ISBN: 978-0-615-20720-9

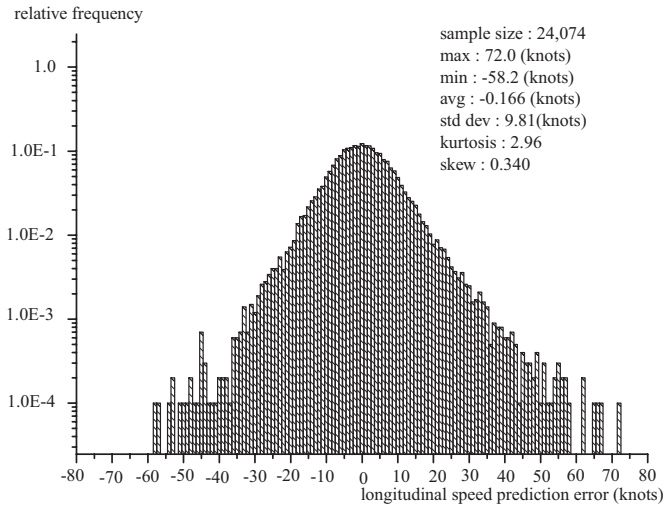

Figure 8. Empirical distribution of longitudinal speed prediction errors (Data omitted, NOPAC)

IV. DISTRIBUTION MODEL

On first sight, the empirical distribution given in Fig. 8 follows a normal distribution. Fig 9 shows the QQ-plot (quantile-quantile plot) for normal distribution. If the empirical distribution follows a normal distribution, the dots in Fig. 9 are on the red straight line. When 'sample quantiles' is larger than -20 and smaller than 20, the dots seem to be on the straight line, however, it is not the case for the data set outside of [-20, 20].

The average and standard deviation of restriction of the empirical distribution on [-20,20] are -0.1142 and 7.757, respectively. The histogram in Fig. 10 shows the empirical distribution and the graph of the probability density function of the normal distribution with average = -0.1142 and standard deviation = 7.757. Fig.10 shows that this normal distribution fits the empirical distribution well.

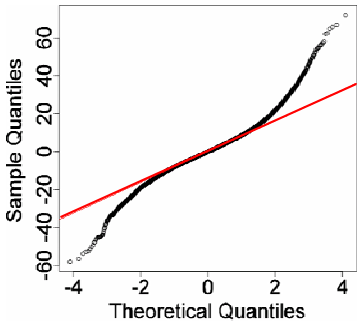

Figure 9. QQ-Plot of Fig.8 for normal distribution

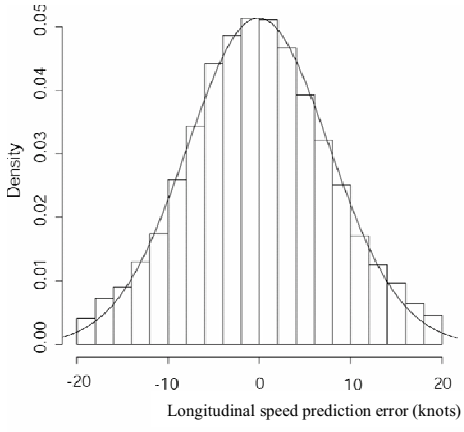

Figure 10. QQ-Plot of Fig.8 for normal distribution

The author applied POT technique to the data set. Extreme value theory claims that the conditional probability Pr{Y<y|Y>u} of distributions satisfying some technical assumptions approximately follows a generalized Pareto distribution when u is large enough. (More precisely, for any distribution which is in the domain of attraction, Pr{Y<y|Y>u} weakly converge to the generalized Pareto distributions as u $\to\infty$.) The cumulative distribution function of a generalized Pareto distribution is given by

$$H(y) = 1 - \left(1 + \xi \frac{y}{\sigma}\right)^{-1/\xi}, \quad 1 + \xi y/\sigma > 0 \cdot \qquad (12)$$

When the shape parameter $\xi<0$, the generalized Pareto distributions are Beta distributions. ($0<y<-\sigma/\xi$) In $\xi=0$, they are exponential distributions and they are Pareto distribution in the case where $\xi>0$. (See [4] and other related papers for more detail.)

The author analyzed the both-side tails of Fig. 8 using POT technique. The R-package extRemes [8] is utilized for the analysis. (R is a free statistical software for data analysis.) For the right tail, we set the threshold u = 20 considering the stability of estimated shape parameter ξ and scale parameter σ. The number of excesses of thresholds is 687 (2.85% of the whole data set). By maximum likelihood method, we found ξ = 0.0426 and 95% confidence interval is [-0.03419, 0.13204]. σ = 7.63 and its standard error is 0.4342. Fig. 11 and Fig. 12 show the QQ-plot and the density plot of this model, respectively. Since almost all dots are on the diagonal line in the QQ-plot diagram, the generalized Pareto distribution fits the right tail of the empirical distribution well. The fact ξ=0.0426 suggests that the right tail is slightly thicker than an exponential distribution.

ISBN: 978-0-615-20720-9

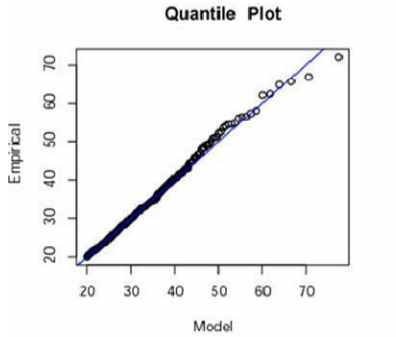

Figure 11. QQ-Plot of generalized Pareto distribution for the right tail of Fig. 8

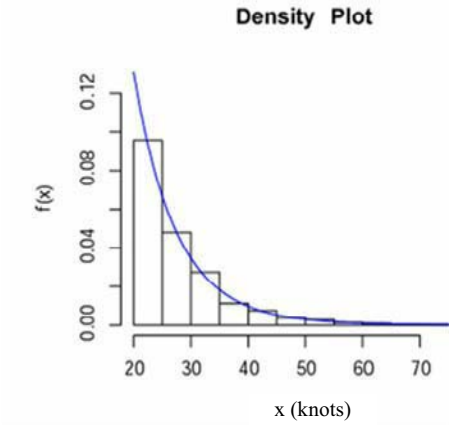

Figure 12. Density plot of generalized Pareto distribution for the right tail of Fig. 8

The same analysis was conducted for the left tail. We set the threshold u = 34. The number of excesses of thresholds is 47 (0.195%). ξ= -0.4484 with 95% confidence interval [-0.61369, -0.08794]. σ=12.22 and its standard error is 2.374. Fig. 13 and Fig. 14 show the QQ-plot and the density plot of this model, respectively. Because of small amount of data set, the estimated parameter has large standard deviation and some dots are apart from the diagonal line in QQ-plot diagram. Hence we cannot apply POT technique to determine the shape of left tail. One way to find the shape of right tail is to assume that both left and right tails follow the exactly same distribution.

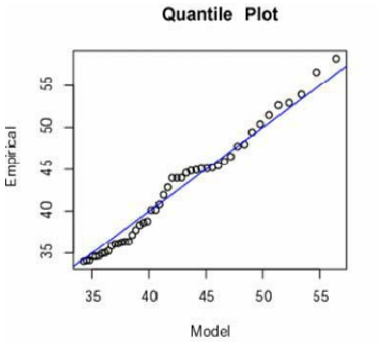

Figure 13. QQ-Plot of generalized Pareto distribution for the left tail of Fig. 8

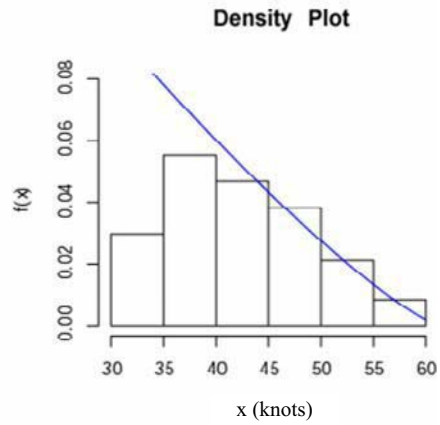

Figure 14. Density plot of generalized Pareto distribution for the left tail of Fig. 8

Under the assumption that both left and right tails follow the same distribution, we applied POT technique to find the shape of the tail. In this case, we have only to analyze the tail of the absolute value of empirical data. We set the threshold u = 20. The number of excesses of thresholds is 1161 (5.26%). ξ= 0.0386 with 95% confidence interval [-0.01962, 0.10373]. σ = 7.093 and its standard error is 0.3048. Fig. 15 and Fig. 16 show the QQ-plot and the density plot of this model, respectively. The generalized Pareto distribution fits the tail of the empirical distribution well judging from QQ-plot diagram.

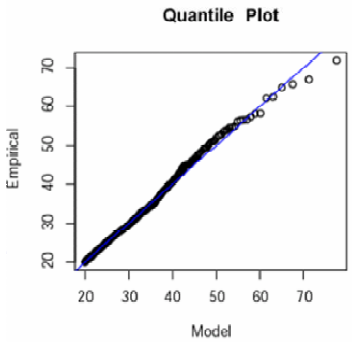

Figure 15. QQ-Plot of generalized Pareto distribution for the both tails of Fig. 8

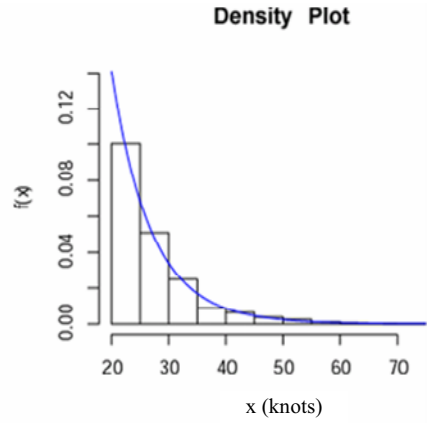

Figure 16. Density plot of generalized Pareto distribution for the both tails

ISBN: 978-0-615-20720-9

of Fig. 8

Let Φ(x) be the standard normal cumulative distribution function, namely,

$$\Phi(x) = \int_{-\infty}^{x} \frac{\exp(-u^2/2)}{\sqrt{2\pi}} du \cdot \quad (13)$$

Then the cumulative distribution function F(x) of longitudinal speed prediction errors, x in knots, is given by

$$F(x) = \begin{cases} \frac{5.26/2}{100} \times \left(1 + 0.0386 \frac{-x-20}{7.093}\right)^{-1/0.0386} & x \le -20 \\ \frac{2.63}{100} + \frac{94.74}{100} \times \frac{\Phi(\frac{x-(-0.1142)}{7.757}) - \Phi(\frac{-20-(-0.1142)}{7.757})}{\Phi(\frac{20-(-0.1142)}{7.757}) - \Phi(\frac{-20-(-0.1142)}{7.757})} & -20 < x < 20 \\ \left(1 - \frac{5.26/2}{100}\right) + \frac{5.26/2}{100} \times \left[1 - \left(1 + 0.0386 \frac{x-20}{7.093}\right)^{-1/0.0386}\right] & x \ge 20 \end{cases} \quad (14)$$

The simplified form is given by

$$F(x) = \begin{cases} 2.63 \times 10^{-2} \times \left(1 - 5.44 \times 10^{-3} \times (x+20)\right)^{-25.9067} & x \le -20 \\ 2.63 \times 10^{-2} + 0.9569 \times (\Phi(\frac{x-(-0.1142)}{7.757}) - 5.18 \times 10^{-3}) & -20 < x < 20 \\ 0.9737 + 2.63 \times 10^{-2} \times \left(1 - \left(1 + 5.44 \times 10^{-3} \times (x-20)\right)^{-25.9067}\right) & x \ge 20 \end{cases} \quad (15)$$

The analysis of both-side tails suggests that the tail of the empirical distribution of longitudinal speed prediction errors follows a Pareto distribution which has slightly thicker tails than exponential distributions. It is hard to calculate the longitudinal overlap probability if the tail follows a Pareto distribution. Even if the tail of empirical distribution is thicker than the exponential distribution, the shape parameter ξ is so small that we may assume that the tail follows an exponential distribution in many cases.

We can assume that |longitudinal speed prediction error|$_{[20,\infty)}$-20 follows an exponential distribution whose probability density function is exp(-x/λ)/λ. Here |longitudinal speed prediction error|$_{[20,\infty)}$ denotes the restriction of the absolute value of longitudinal speed prediction error on [20,∞). The maximum likelihood estimator of λ is the average of the empirical data set of |longitudinal speed prediction error|$_{[20,\infty)}$-20. It is 7.439. Hence we get the following probability density function of a model of longitudinal speed prediction errors.

$$f(x) = \begin{cases} \left(1 - \frac{5.26}{100}\right) \times \frac{\frac{\exp(-(x+0.1142)^2/(2*7.757^2))}{\sqrt{2\pi}*7.757}}{\Phi(\frac{20-(-0.1142)}{7.757}) - \Phi(\frac{-20-(-0.1142)}{7.757})} & |x| < 20 \\ 2.63 \times 10^{-2} \times \frac{\exp(-(|x|-20)/7.439)}{7.439} = & |x| \ge 20 \end{cases} \quad (16)$$

V. CONCLUSION

This paper first reviews the basic concept of ADS-C and summarizes the methodology to find the longitudinal speed prediction errors of ADS-C.

A few ADS-C reports, which are not coupled with the fixed intent group and the estimated time of arrival at the next waypoint is close to the stamped time, have large longitudinal prediction errors in magnitude. (Max. 668 knots)

Fig. 8 shows the empirical distribution of the longitudinal speed prediction errors of periodic report pairs which are coupled with both fixed projected intent group and predicted route group and whose reporting time interval is greater than 26 min and less than or equal to 27 min. By QQ-plot, it turns out that this distribution on [-20, 20] follows the normal distribution whose average is -0.1142 and whose standard deviation is 7.757. POT (Peak over Threshold) technique of Extreme Value Theory was applied to find the shape of the tail of Fig. 8. The tails (outside -20 and 20) follow the generalized Pareto distribution whose shape parameter ξ= 0.0386 and the scale parameter σ = 7.093. Equation (15) gives the explicit description of the cumulative distribution function.

The density function of the distribution in (16) is also given under the assumption that the tails follow an exponential distribution.

ACKNOWLEDGMENT

The author thanks Japan Civil Aviation Bureau, ATM Center, Kobe Aeronautical Satellite Center and Mr. Sasamoto for the provision of information and data sets.

REFERENCES

[1] Air Traffic Service, Annex 11 to Convention on International Civil Aviation Section 2.26

[2] Anderson, D. A Collision Risk Model Based On Reliability Theory That Allows For Unequal Navigational Accuracy, ICAO SASP-WG/WHL/7 WP/20 REVISED, May 2005.

[3] Fujita, M. Nagaoka, S and Amai, O. Safety Assessment prior to Implementation of 50NM Longitudinal Separation Minimum in R220 and R580, ICAO SASP-WG/WHL/9 WP/14, May 2006.

[4] Coles, S. An Introduction to Statistical Modeling of Extreme Values, Springer Verlag London, 2001.

[5] Automatic Dependent Surveillance (ADS), ARINC Characteristic 745-2, June, 1993.

[6] Minimum Operational Performance Standards for Airborne Automatic Dependent Surveillance (ADS) Equipment, RTCA DO-212 October, 1992.

[7] ATS Data Link Applications over ACARS Air-Ground Network, ARINC Characteristic 622-4, October, 2001.

[8] http://www.assessment.ucar.edu/toolkit/

ISBN: 978-0-615-20720-9