RNP REQUIREMENTS FOR 4-D NAVIGATION

Ir. P.B. Ober, Dr. ir. E. Theunissen, Ir. R.C. Meijer and Prof. dr. ir. D. van Willigen
Delft University of Technology, Faculty of Information Technology and Systems

BIOGRAPHY

Bastiaan Ober's areas of experience include the influence of multipath on GPS positioning, carrier phase differential GPS, ambiguity resolution and integrity monitoring. He is currently working as a Ph.D. student doing research on integrity monitoring algorithms for integrated navigation systems.

Erik Theunissen heads the DELPHINS II program (research into spatially integrated data-presentation for aircraft navigation and guidance) sponsored by the Dutch Technology Foundation STW. DELPHINS II includes the use of digital terrain data for terrain depiction.

Marco Meijer has been working in the field of radio navigation and aviation on various topics such as integrated navigation systems, datalinks, propagation modelling, software engineering and air traffic control.

Durk van Willigen is appointed professor in radio-navigation at the Delft University of Technology and president of Reelektronika b.v., a small consultancy company in the field of radar and radionavigation. The International Loran-C Association awarded Dr. Van Willigen the Medal of Merit in 1996.

ABSTRACT

This paper discusses possibilities to extend the current required navigation performance concept that does not cover 4-D navigation. It focuses on the way the threshold to generate tunnel incident alarms should be chosen. A dynamically varying threshold is proposed to provide the required safety by generating timely alarms while keeping the amount of false alarms as small as possible.

1. INTRODUCTION

Required Navigation Performance (RNP) is a concept designed to achieve a desired Target Level of Safety (TLS) [Kelly94]. It is currently only accepted and specified for the lateral part of the navigation process (LNAV) of the en-route phase [ICAO94]. Vertical navigation (VNAV) boundaries have not been defined yet, nor is RNP defined for approach and landing. However, proposals for RNP precision approach and landing available [AWOP][MASPS]. All these proposals give suggestions for straight approaches only.

This paper will address the alerting mechanism which is needed and a parameter that is not a part of the original RNP concept: time. The main discussion focuses on the importance of time in relation to the tunnel incident alarm that the RNP concept defines. While concentrating on lateral navigation and flying straight segments, the notions that are introduced should be helpful for extending the concept to vertical and non-straight segments as well and can be seen as a first step towards the definition of a full 4-D navigation concept.

One of the reasons for writing this paper lies in the need for the time dimension in the integrity research at Delft University. Therefore, special attention is paid to integrity monitoring in relation to time. The detection capability of the pilot is addressed as well. The likelihood of timely detection improves with better situational awareness, that can be provided by advanced displays.

2. THE RNP CONCEPT

RNP defines four performance parameters. The desired TLS is accomplished by distributing the allowed probability that a specified tunnel in space is left (tunnel incident) among these four parameters. First of all, *accuracy* is measured by the Total System Error (TSE), that should remain within the tunnel specifications. The TSE is the sum of the Flight Technical Error (FTE), that can be measured, and the Navigation Sensor Error (NSE) that can be estimated. *Continuity* measures the risk that guidance is lost, while *integrity* ensures that the pilot is notified of such loss. Finally, *availability* defines the probability that system guidance is indeed present.

RNP assumes that an aircraft is established on a stable flight path. Based on the dynamic response of the aircraft and its navigation system, the flown track can then be divided into sections with sufficiently uncorrelated positions. The aircraft has the opportunity to leave the tunnel during each of these segments due to excessive flight technical or navigation system errors. RNP is based on the ability to determine the likelihood of such events. In addition, the original RNP concept [Kelly94] proposes a mechanism called *tunnel incident alarm* to decrease this likelihood by making pilots aware of (impending) insufficient system performance. This enables them to take appropriate action in case of a loss of guidance or deteriorated performance. In the RNP MASPS [MASPS, p.53-54] the alerting mechanism is listed as optional. The tunnel incident alarm will be discussed in more detail in section 3. First, some attention will be directed to problems that have to be solved when extending the current RNP concept to full 4-D navigation.

2.1 Extension to 4-D Navigation

The original RNP concept [Kelly94] clearly focused on the lateral part of straight-in approaches. The proposed method has served as a basis for the RNP specifications of precision approach and landing in [DO226]. In [AWOP] it is stated that "*Once the tunnel for the straight-in final approaches matures, attention can be directed toward curved approach tunnels where procedures will include 3D and 4D RNAV*". Furthermore, [ICAO94, $3.2.3] states that "*No consideration is currently given to time or vertical navigation for the purpose of establishing RNP types for en-route operations*".

Current RNP does therefore not yet cover 4-D navigation. Still, when aircraft fly procedures described in terms of time intervals, the time dimension is very important, both for the aircraft and air traffic control (ATC). Aircraft also use time, sometimes in relation to distance, to determine flap settings or more general configuration changes. ATC benefits from aircraft using time-to-arrival control to minimise separations and optimise runway utilisation.

As compared to the original RNP concept, 4-D navigation requires a full 3-D position rather than only a 1-D cross-track position constraint. Furthermore, the position constraint is going to be varying as a function of time. One might expect that the extension of RNP to a time-varying containment area involves nothing but the translation of the time-requirement into a desired position, and that 3-D position deviations can be treated just as the cross track TSE was treated. However, additional considerations will complicate this process, many of which will be related to notions like 'time' or 'speed'. An example is provided in the next paragraph.

2.1.1 Along track errors

A good example of extra requirements that might play a role when extending the current RNP concept is the relation between the along track containment area and the economically optimal airspeed at which aircraft are preferably operated.

The 4-D position requirements can be translated into allowed margins for the reference groundspeed of an aircraft from which a reference airspeed can be derived using wind-estimates. An error in the wind-estimates used during planning leads to a difference between the optimal and reference airspeed. It can be expected that the spatial margins that will be allowed are going to be partly exploited to operate the aircraft at the optimal airspeed, which will influence the distributions of the along track position deviation. This distribution will also be influenced considerably by the accuracy of the wind estimates. This should be taken into account when defining an alerting mechanism.

3. THE TUNNEL INCIDENT ALARM

To reduce the likelihood of an excessive total system error (TSE), RNP has introduced the *tunnel incident alarm*. The tunnel incident alarm is a warning that notifies the pilot when the aircraft is about to leave the tunnel due to:

- fault-free excessive TSE
- aircraft failures
- piloting failures
- environment phenomena

[Kelly94] describes a way of setting the alarm threshold, based on a snapshot measurement of the FTE and a corresponding standard deviation. This method will therefore only work as long as:

- A snapshot of the current TSE is a reliable indication of the future TSE
- The rate of error build-up is sufficiently low

Here 'sufficiently low' implies that the pilot has enough time to react adequately to the alarm and maintain the target level of safety. En route, this means that the pilot can keep the aircraft inside the tunnel. During an approach, the time should be sufficient to make a safe transition to a missed approach procedure.

In general, the error build-up will depend on the situation and the kind of failure. A high position error rate build-up can occur when flying curved segments or transitions between straight and curved segments. In the latter situation, the rate at which the FTE can increase is high when the curve is not initiated well. Consequently, the current TSE provides a less reliable indication of the future TSE. Moreover, the tracking performance in curved segments is lower than the performance in straight segments. When the outer tunnel boundary is fixed, the threshold for the tunnel incident alarm should therefore become smaller to allow for the increased uncertainty in the future TSE to maintain the same TLS. This, however, will increase the alarm rate.

Because the underlying cause is the less reliable estimate of the future TSE, a potential solution to guarantee timely warnings without unnecessarily increasing the false alarm rate would be to improve this estimate by incorporating knowledge on the error build-up rate. The problem will be the same for along track, cross track and vertical deviations, although the error distributions involved will differ. Although the following discussion will focus on cross track errors, there is no conceptual difference with the handling of along track or vertical errors, making the discussion relevant for 4-D navigation as well.

Explicitly using the error build-up rate leads to a dynamic TSE threshold that becomes smaller for higher build-up rates. Determination of this threshold can be based on the (estimated) time until the edge of the containment volume is reached when no course corrections would be applied. This leads to the approach that will be discussed in more detail in the next section.

4. THE TSE THRESHOLD

The introduction of a tunnel incident alarm is only useful when there is sufficient time to react adequately. The minimum amount of required correction time therefore seems a very natural starting point for setting alarm thresholds. One could say that a tunnel incident alarm should be issued *before the aircraft passes the point from which staying inside the tunnel becomes impossible*. The location of this point depends on the needed reaction time, the velocity of the aircraft and the direction of flight.

[Kelly94] introduces an alarm based on a predetermined threshold. When only a snapshot of the TSE is used to decide on an alarm, this necessarily implies that the TSE threshold is static as the decision to generate an alarm can be based on the current position only. When such a static TSE threshold is to provide the pilot with a certain fixed amount of time to allow sufficient corrections *in all situations*, its magnitude should be selected on a worst case basis. First of all this requires an assumption to be made on the maximum track angle error and velocity.

Using a worst case, however, is overly conservative and will increase the false alarm rate. An alternative is to define a *dynamic* TSE threshold, depending on aircraft velocity and direction of flight, that guarantees a certain minimum separation towards the boundary of the outer tunnel. Two approaches will be discussed that would provide such a threshold when tracking a straight segment. In the discussion, the following assumptions have implicitly been made:

- On a straight segment, the pilot is keeping lateral acceleration zero. The future position of the aircraft can be adequately predicted with a first order model.
- When pilots apply corrective action after a tunnel incident alarm, they fly a curved segment with a constant bank angle. The future position of the aircraft

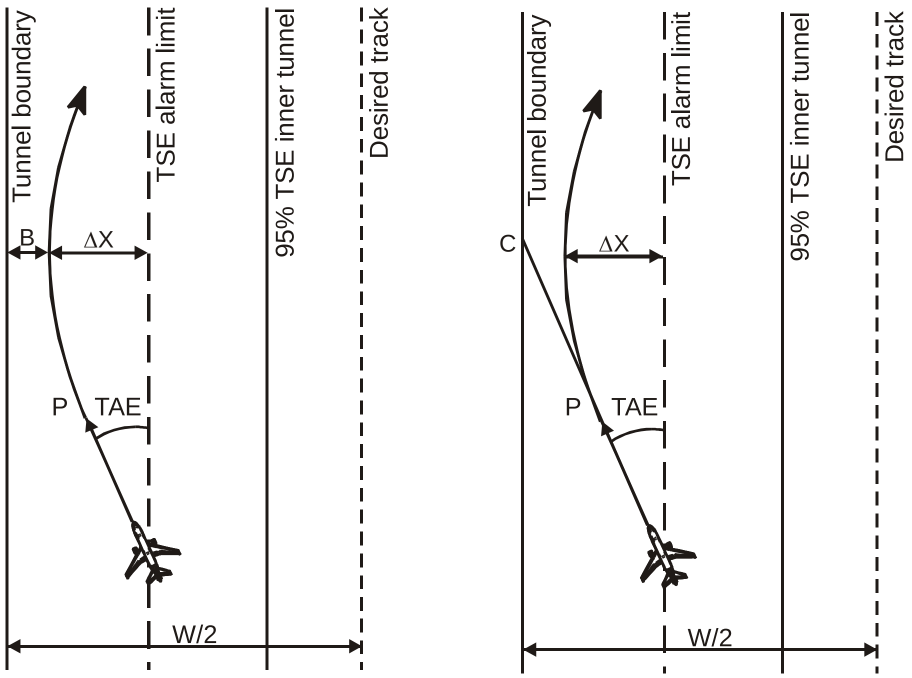

Figure 1a&b. *Two ways to determine the tunnel incident alarm TSE threshold: by using a fixed spatial buffer or a fixed temporal margin.*

can be adequately predicted with a second order model.

- An alarm is only needed when the information that something is wrong is not dominantly available from the instruments needed to accomplish the tracking task.

4.1 Causes for FTE

When tracking a straight segment, the reference bank angle is zero. During manual control, a deviation from this bank angle should alert the pilot that something is wrong. It can be assumed unlikely that a pilot will not notice a deviation from zero bank angle[i]. Nevertheless the aircraft may still fly a faulty track. For supervisory control, incidents have occurred caused by an incorrect aircraft track in combination with a monitoring failure of the pilots which would not have occurred when the pilots where manually controlling the aircraft.

A possible cause for an unnoticed position error build-up is when a pilot flies a pre-computed heading which as a result of a change in conditions yields an incorrect track. With today's instruments a track is often flown based on a certain heading that is derived from the desired track, the wind direction and velocity estimates. Figure 2a illustrates this situation. The ability of the pilot to detect that the computed heading might no longer be appropriate to follow the desired track depends on the presented data and the way it is displayed. This will be further discussed in section 6.

When the track of the aircraft does not coincide with the desired track, the position error will increase with time at a rate that is proportional to the magnitude of the track angle error. When the position error exceeds the TSE alarm limit and an alarm is generated, the pilot will have to use corrective action to bring the aircraft back the desired track. However, the position error will still be growing until the aircraft direction has been adjusted sufficiently. The maximum position error that results will be a function of track angle error, velocity, reaction time, and the magnitude of the corrective action. Under the assumption that the pilot reacts as fast as possible and will apply a control action that will return the aircraft to the desired track as fast as possible, the maximum position error will vary as a function of:

- velocity
- track angle error
- maximum bank angle that will be applied

4.1.1 Fixed temporal margin

One of the ways to determine a TSE threshold could be to use a fixed temporal margin after which the aircraft would reach the tunnel boundary (point *C* in Figure 1b) when no action would be taken. This temporal margin should be the sum of a minimum reaction time T_R and the time needed to correct the course of the aircraft T_C. Since the maximum position error after corrective action is a function of velocity, track angle error, and the maximum bank angle the pilot will allow in the turn, T_C should be selected on a worst case situation.

4.1.2 Fixed spatial buffer

A natural extension of the previous approach is to *compute* T_C rather than use a worst case situation. In this case an easy way to set the threshold can be based on an estimate of the additional position deviation ΔX which occurs after the pilot applies a corrective action rather than on T_C explicitly. Figure 1a illustrates this concept. Under the assumption of applying a fixed bank angle, this additional deviation can be approximated by:

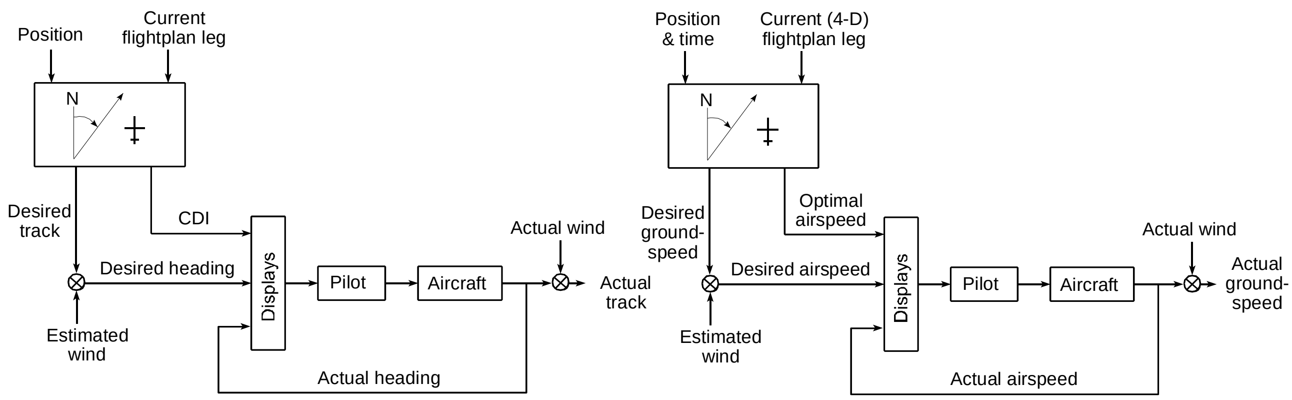

Figure 2a&b. *Left hand side: when the actual heading differs from the desired heading, the ability of the pilot to detect this depends on the displayed data and the way of presentation. Right hand side: To allow aircraft to fly 4-D routes at the most efficient airspeed, an estimate of the expected wind direction and velocity along the route to be flown must be used to calculate the time constraints. When during the execution of the 4-D route the actual wind differs from the estimate used in the definition of the route, the reference airspeed needed to stay in the centre of the 4-D containment area will differ from the optimal airspeed. Fortunately, aircraft do not have to fly in the center, but only within certain margins around it. As long as the optimal airspeed allows the aircraft to stay within the longitudinal constraints there is no need to deviate from it.*

$$\Delta X = V \cdot T_R \cdot \sin(TAE) + \frac{V^2(1 - \cos(TAE))}{g \cdot \tan(\phi)}$$

in which

V: the velocity of the aircraft
T_R: the reaction time required until the pilot starts the turn
TAE: the track angle error at the moment the TSE threshold is exceeded
ϕ: the bank angle in the corrective turn.

By defining a fixed spatial buffer B, the TSE position error threshold TSE_{ALARM} follows from:

$$TSE_{ALARM} = \frac{W}{2} - B - \Delta X$$

in which W is the tunnel width.

Note that this approximation neglects the transition from the straight segment to the curved segment. By imposing a limit on the maximum bank angle in the turn, an estimate of the minimum additional deviation is possible.

The two approaches for choice of a TSE threshold are quite similar. The fixed spatial buffer concept could be used to calculate a minimum required T_R for the fixed temporal margin concept, based on the specification of a worst case velocity V and track angle error TAE, and maximum bank angle ϕ.

4.1.3 Discussion

As is indicated in [Kelly94], the TSE alarm tunnel lies between the inner and the outer tunnel. The magnitude of the fixed TSE threshold should be selected so, as to allow correction in such a way that the aircraft does not leave the containment area. As has been discussed earlier, the possibility of the aircraft to stay inside the containment area is a function of velocity and track angle error at the moment the alarm is generated and the maximum bank angle that will be used in the corrective manoeuvre. The use of a fixed TSE threshold to warrant that the aircraft does not leave the containment area implies that an assumption has been made regarding the maximum magnitude of these parameters. Beyond a certain velocity, a minimum track angle error will exist beyond which the aircraft will cross the outer tunnel regardless of the corrective action. If this can occur at a realistic velocity, while the cross track error is smaller than the RNP parameter, no useful TSE alarm limit can be defined. In [Clarke98] the need for tighter RNP requirements is identified as one of the key factors to enable short final captures. Such an increase will reduce the minimum track angle that would allow for timely recovery with a fixed TSE threshold.

The only alternative to provide a timely warning is a mechanism which takes the magnitude of the track angle error into account in the decision to generate an alarm. Both previously discussed methods can be applied for this purpose.

5. INTEGRITY MONITORING

As discussed above, an alarm should be supplied to the pilot when he is about to leave the tunnel. This alarm should be given once the position error exceeds a certain threshold that depends on the time constraints for an adequate reaction. Unfortunately, the position error is only approximately known as the position information that the navigation system provides always contains uncertainty. Therefore, the above statement has necessarily to be relaxed. In reality, *an alarm should be issued as soon as the position error is sufficiently likely to lie outside the TSE alarm limit,* where the TSE alarm limit is the limit discussed above. The tunnel incident alarm has been proposed to be obtained by uplinking an estimated NSE from the ground [AWOP]. It is also possible however, to let the onboard navigation system generate the alarm itself. To be able to do this reliably, it is important to assess the navigation system performance.

When the navigation system functions according to specification, the errors will have well defined properties and will be small compared to the dimensions of the tunnel. However, when there is some kind of failure, there will be an extra error in the position determination. The situation is sketched in Figure 3. The smaller ellipsoid is the area that contains 95% of the position errors, and represents the navigation system performance in the absence of failures. The presence of system failures will cause a much larger variation of position errors. Integrity monitoring can be used to reduce the tails of the distribution of the Navigation Sensor Error (NSE) and control the size of the outer ellipsoid, that contains the actual position with a probability near to one.

Under the condition that there is sufficient navigation system integrity available, substantial position errors can be detected by the integrity monitoring functions in the navigation system. Integrity is obtained by checking for mutual consistency among redundant position information. When there is so much inconsistency that the probability of being outside the TSE limits become unacceptably high, the integrity check should result in an *failure detection*. Sometimes, the erroneous measurement(s) can be removed from the position solution and navigation can continue. When this is not possible, or when insufficient usable measurements remain, the error detection should be followed by an alarm to warn the pilot that the position information contains errors. Moreover, when the redundancy is insufficient to provide sufficient failure detection capability, a warning should be issued that the integrity of the position information can not be guaranteed.

Time is an important parameter in integrity monitoring, as

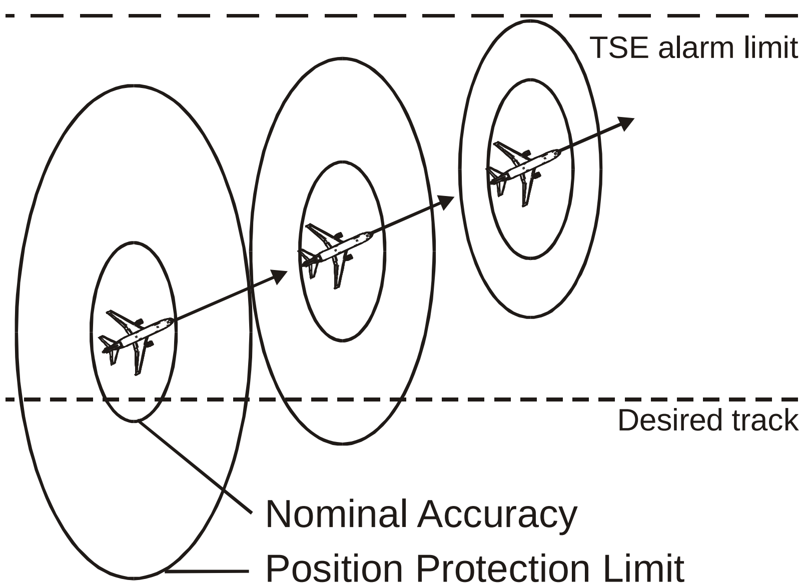

Figure 3. *The navigation system can guarantee the position within the position protection limit. The more time there is available to detect large system deviations caused by sensor failures, the better the protection gets.*

the containment volume that a navigation system can guarantee (the outer ellipse in the figure) depends on it. The more time there is available for error detection, the more information the system can collect to support the decision on the detection of an error. As a result, the tails of the position error distribution are reduced more effectively when there is more decision time available. This implies that the outer ellipsoid in the Figure 3 will become smaller and the number of false alarms will decrease.

Unfortunately, it is not easy to determine the integrity monitor's available decision time. Note that 'decision time' here means*: the time over which already recorded measurements can be used in the position error detection scheme*, rather than referring to the future as was the case in the TSE prediction! One of the reasons why determining the decision time is so hard, is that in case of a past or current system failure that is yet to be detected, *the current and future TSE depend on the past behaviour of the failing sensor*.

The following two mechanisms are important.

1. When a failure has occurred at some unknown time t_F somewhere in the past, using uncorrupted measurements from before t_F would reduce the effectiveness of the error detection scheme rather than enhance it. On the other hand, including more measurements can help to reveal smaller errors that can not be detected using too limited an amount of samples.

2. The aircraft can act as a low-pass filter in translating sensor errors to position errors: fast increasing (undetected) failures will not immediately influence the position dramatically, due to the inertia of the aircraft.

When the failure behaviour of the sensor is unknown, it should be possible – at least in principle – to compute some 'worst case behaviour', including a 'worst case t_F' and error profile. Here worst case would mean: giving the largest probability that at the current moment an alarm should be issued.

Special care should be taken when errors in the navigation system might lead to erroneous direction information that is used in the TSE threshold determination. Whether this is indeed the case depends on the way this information is derived: are independent instruments used, or is the position information from the system itself differentiated? If both the position and the flight direction – possibly even speed – are derived from the same corrupted data extreme care should be taken in the determination of the optimal TSE threshold and alarm generation threshold.

The problem of determining feasible ways to arrive at generally applicable thresholds that take all important features into account seems extremely complicated. More research is definitely necessary!

6. RELATION WITH DATA PRESENTATION

Besides an alerting mechanism, the likelihood of a missed detection can also be reduced by data presentation concepts that increase the pilot's navigational awareness. The way data is presented determines the mental effort needed for interpretation and evaluation, and the likelihood that

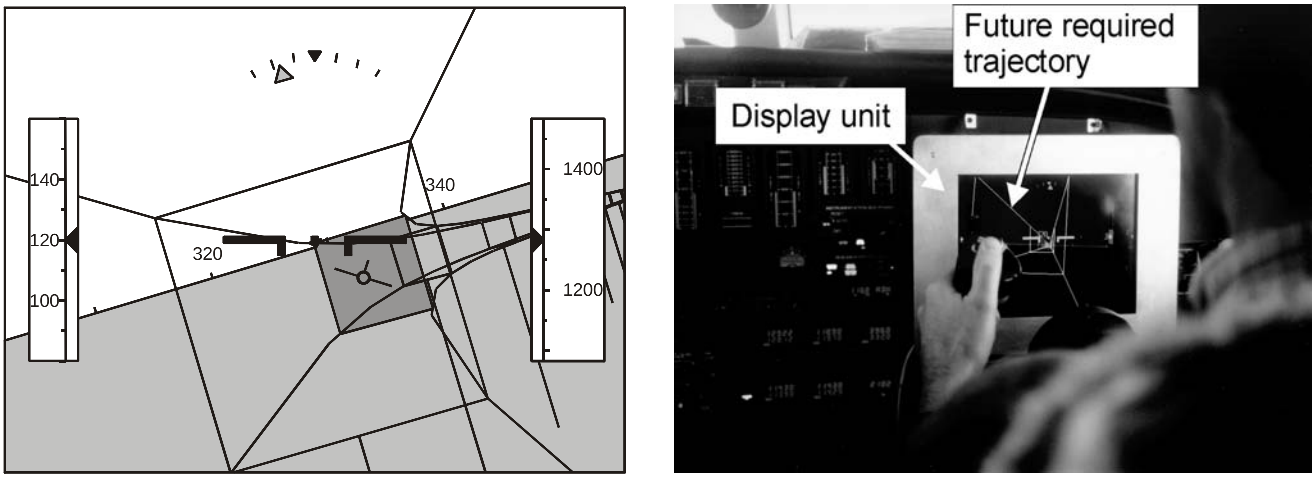

Figure 4a & b. *The perspective flightpath display developed by Delft University*.

deviations from the desired position and/or course are detected timely. In [MASPS] some general user interface requirements are presented but these still allow for a wide range of implementations.

Instruments which directly show the desired track relative to the current earth-referenced direction of flight should reduce the likelihood that a growing position error goes undetected. Today’s instruments often provide a separate indication of the various variables which must be monitored by the pilot. When integrating data into these instruments to allow the pilot to determine whether all spatial constraints are satisfied, the altimeter tape could include a depiction of the allowed vertical margins, the navigation display could depict the allowed horizontal margins, and the speed tape could include a depiction of the aircraft referenced velocity constraints following from the fourth dimension, time. To establish sufficient spatial and navigational awareness, pilots have to scan the various instruments and mentally integrated the obtained information. The additional data integrated into these instruments will increase the dwell time and the task demanding load. Several new display concepts are being explored which aim to reduce the task demanding load by providing integrated 3-D and 4-D navigation information. With so-called perspective flightpath displays (Figure 4a) the required mental integration of spatial data is made superfluous by providing a spatially integrated presentation of the desired trajectory and its constraints. The magnitude of the constraints visualised by perspective flightpath displays is not necessarily an indication of the tunnel boundary or other RNP parameters, but is typically based on tracking performance and control activity considerations [Theunissen97]. In Europe, both Delft and Munich University of Technology have developed, implemented and flight tested these types 4-D navigation and guidance displays. Figure 4b shows the display during an in-flight test which was performed by Delft University in 1994.

7. CONCLUDING REMARKS

When extending the RNP concept to 4-D, the distribution of the along track position errors may show a shift from the center of the desired position. This must be taken into account when defining TSE alerting mechanisms. RTCA RNP MASPS [MASPS] allow for a range of navigation and guidance data presentation methods. The likelihood of a timely detection of a situation in which the aircraft is about to leave the RNP outer tunnel is related to the way guidance and navigation data is presented. The likelihood of a missed detection of a potentially dangerous situation can be reduced through the introduction of an alerting mechanism. In the [Kelly94] a fixed TSE threshold alerting is proposed. When operating in RNP airspace, this mechanism aims to provide the pilot with a timely warning. In this paper it was pointed out that the selection of the TSE alarm threshold requires a tradeoff to be made between the amount of events in which a timely alarm is generated and the amount of false alarms. Furthermore, it was illustrated that for smaller RNP values the fixed TSE concept may not be adequate. Two mechanisms, one based on a fixed temporal margin and one based on a fixed spatial buffer, have been proposed as alternatives which can provide timely warnings. Besides an alerting mechanism, the likelihood of a missed detection can also be reduced by data presentation concepts which increase the pilot’s navigational awareness. An example is the tunnel-in-the-sky display format which has been developed and flight tested by Delft University of Technology.

REFERENCES

[AWOP] “RNP Tunnel concept for precision approach and landing”, information paper for the All Wheather Operations Panel , 14th meeting, Montreal, 1993

[Clarke98] J.-P. Clarke, J.-P, R. K. Nicholson, A. Haraldsdottir and A. W. Houck, “Simulator Evaluation of

Autopilot-Coupled Short Final Approaches", Proc. of the AIAA Guidance Navigation and Control Conference, Boston, MA, August 10-12, 1998

[DO226] RTCA DO-226 Guidance Material for Evolving Airborne Precision Area Navigation Equipment with Emphasis on MLS, issued 5-25-95

[ICAO94] ICAO Doc 9613-AN/937; "Manual on required navigation performance (RNP)", first edition, 1994

[Kelly94] Kelly, R.J. and J.M. Davis; "Required Navigation Performance (RNP) for Precision Approach and Landing with GNSS Application", Journal of the Institute of Navigation, Vol. 41, No. 1, pp. 1-30, 1994

[MASPS] RTCA DO-236, Minimum Aviation System Performance Standards: Required Navigation Performance for Area Navigation, issued 1-27-97

[Theunissen97] E.Theunissen; "Integrated Design of a Man-Machine Interface for 4-D Navigation", Delft University Press, 1997

---

[i] Still, incidents have occurred in which the autopilot disengaged and the aircraft attained a significant bank angle before the pilots detected this and applied corrective action.