A98-31689 ICAS-98-6,7,2

# EVALUATION OF WINDSHEAR HAZARD DISPLAYS AND GO-AROUND PROCEDURES USING PILOTED SIMULATIONS AT NLR.

by
H. Haverdings & W.F.J.A. Rouwhorst
National Aerospace Laboratory NLR
P.O. Box 90502,
1006 BM Amsterdam, The Netherlands

Abstract

Two piloted simulation experiments were conducted by NLR, in 1994 and 1996, to evaluate a number of novel windshear icon displays, driven either by a scanning laser (1994) or by an Airborne Doppler (weather) Radar (ADR) (1996). In addition several flight procedures of coping with the windshear threat information were also evaluated, in terms of speed additives (1994) or aircraft configuration changes (1996). In the experiment of 1994 also the concept of speed feedback was evaluated: this concept raised the alert thresholds used to warn the crew of windshear whenever the pilot increased the approach speed set in the Autothrottle. This speed feedback concept, although initially appreciated much by the pilots, was later denounced due to its tendency to induce crews to penetrate the windshear. The windshear icon displays evaluated were liked very much by the pilots. Especially the ADR was favored because of its "natural" interface in the cockpit, allowing pilots to correlate weather information with windshear information. Its drawback of not being useful in case of "dry" microbursts (MBs) becomes less significant the lower the (dBZ) level becomes at which precipitation can still be detected. Integrating the ADR with the reactive sensor was found to be a good protection feature, provided a careful tuning of alert thresholds is done between the two systems. It may even be advisable to not suppress the caution alert from the reactive sensor when combined with the ADR. Overall the icon displays had a positive effect on situational awareness and on flight safety. The crew's workload sometimes increased with the display, however, depending upon the weather situation. In some cases it was the workload of the pilot not flying which increased more than that of the pilot flying, especially during turning escape maneuvers to avoid windshear in the go-around.

## 1 Introduction

At the National Aerospace Laboratory NLR of The Netherlands, a research program had been underway since 1990, regarding the subject of windshear on the approach[1]. Several activities were planned according to a master schedule, in order to prioritize the available resources for the development of models and evaluation of new concepts for detection and display of hazardous weather, especially windshear. The research effort culminated in 3 experiments (in 1993, 1994 and 1996), performed on the NLR's moving-base Research Flight Simulator (RFS), whereby with each experiment new features were added to the previous one. Preliminary, as well as more detailed results of the first and second experiment were reported in ICAS paper 94-7.1.3[1] and ICAS paper no. 96-3.8.3[2]. The final results of the 1994 experiment have been reported by the Group of Aerospace Research technologies EURope (GARTEUR) Action Group FM(AG07)[11]. Reports of the 1993 and 1996 experiments have not (yet) been released in the public domain.

NASA has for a long time been very active in evaluating windshear detection systems. A thorough review of NASA's activities in this area has been described by Arbuckle *et al* [4]. Apart from a prototype $CO_2$ scanning laser, which was installed on-board their B737, NASA also evaluated infra-red sensors in 1992. An integrated evaluation of several sensors, such as a scanning laser or a windshear radar, in conjunction with different flight procedures, had not yet been performed.
These areas were addressed by the research conducted by NLR.
An important aspect of using a forward-looking windshear detection system, in combination with a reactive windshear detection system, is how to cope with the alert information, and what to do in case a particular alert is generated by one system, and not by the other. Much thought was given by NLR about integrating these different types of information, and how to interface them with newly developed, or modified flight procedures. Of particular interest was the so-called Windshear Training Aid[3] (WTA) type of go-around in windshear. One of the goals of the NLR research was to evaluate whether normal go-around procedures could be used instead.

In the 1994 and 1996 simulation experiments novel display concepts were tested to display hazardous windshear. In the experiment of 1994 this was a scanning laser-driven icon display with an additional novel feature called "speed feedback", described later. The severity of the windshears displayed varied from an advisory (color blue) to a warning level (color red). In the last experiment of 1996 an airborne Doppler radar, with an integrated windshear detection mode, was used to detect and display only hazardous windshear.
Description of the experimental objectives and discussion of results is given in the following chapters.

## 2 Experimental objectives

The primary objective in the 1994 experiment was the evaluation of a windshear hazard icon display. In this experiment the unique windshear hazard icon display showed icons of windshear, detected by a scanning laser. Using the Doppler principle the windshear was detected by measuring the relative velocity of the aerosol particles along the laser beam. The detection performance of the laser deteriorated, however, when precipitation was present. Detected wind-shears were displayed in the form of one or more icons, of varying size and color.

Other issues investigated [11], but not discussed here, were the modification of a flight director with a windshear-stall protection mechanism in the go-around, the assistance of the pilot with a flight director or not, and the type of maneuver made in the go-around procedure, i.e. straight ahead or turning away from the threat.

The main objective of the 1996 experiment was to evaluate a windshear hazard display, in conjunction with an Airborne Doppler weather Radar (ADR), together with adjusted go-around procedures. Also here the Doppler principle was used to determine the windshear threat. Contrary to the icon display in the 1994 experiment, windshear threats below the warning threshold were not shown. The display of one or more windshear icons was integrated with the display of weather, obtained from an ordinary weather radar. This was done in order to alleviate the workload in the cockpit, and was also based on pilot comments from the 1994 experiment.

All windshear scenarios contained various rainfall conditions so as to allow the effect of precipitation on the windshear detection performance of the sensors to be studied.

## 3 Windshear models used

The windshear models used were of a variety of types. In the 1994 windshear experiment each weather scenario consisted of one or more microbursts, being located in a particular position (either in the threat position, or not, and of severe strength or not). Each microburst model consisted of a combination of one or more ring-vortices, each one being modeled according to the model of Schultz [5]. In 1994 the FAA, through NASA, released a series of microburst models of varying precipitation intensity, in order to have a certification base available to certify new windshear detection devices developed by industry [13]. Some of these cases were (numerically simulated) catastrophic windshear cases such as the Fort-Worth Dallas crash case. In total a set of 9 such cases was delivered by the FAA on CD-ROM to NLR for implementation in the research flight simulator in 1996.

## 4 Windshear hazard displays used

### 4.1 Laser icon display

The most interesting feature to be investigated consisted of a novel type of ***windshear hazard display***. In the 1994 experiment a scanning laser scanned ahead of the aircraft. Whenever a particular windshear, defined by the so-called F-factor being above certain threshold levels, was detected this data was converted into a size/diameter and intensity of a circular-shaped icon, which was displayed on the navigation display. The diameter of the circular-shaped icon was determined by the peak-to-peak *distance* of the maximum/minimum F-factor [6], while the color of the icon reflected the *intensity* of the hazard, denoted by the factor 'F'. The color displayed could range from blue ($.04<F\leq.10$), to amber ($.10 < F\leq.21$) to red ($F>.21$). An example of such a typical icon display is shown in Figure 1. The value of $F_{crit}=0.21$ was based on the reactive system's 'must alert' boundary for an averaging time of 5 seconds [14].

A special novel feature added to this display was the introduction of ***speed feedback***. If this concept was active then, whenever the pilot would increase the speed regulated by the Autothrottle system by resetting the so-called "bug speed", the critical threshold for F (e.g. .04, .10 or .21 as given above) would be increased proportionally. The concept of speed feedback, and how it interacts with the warning logic and displays, is shown in Figure 2. With this concept it is possible that the color of a displayed icon changes from amber back to blue, for example. In this regard this concept can provide additional information to the crew about the intensity of the windshear.

The drawback of any type of laser is its sensitivity to precipitation: the heavier the precipitation, the less the performance in terms of detection range. This performance degradation was also modeled in the simulation experiment using the lidar equation [7], and could lead to a reduction in icon size, or its disappearance.

### 4.2 Windshear radar display

For the 1996 experiment a so-called ***windshear radar*** model was developed and used. Applying the same Doppler principle as with the laser the radar beam would detect velocity changes of precipitation particles within the radar beam, which could be translated (approximately) into the well-known windshear hazard F-factor. The inertial speed $V$, the gravity $g$ and the range gradient d(..)/dr, operating on the Doppler speed $V_{Doppler}$, are combined to calculate the approximated F-factor, also known as the "radar-F", as follows:

$$F(r) = \frac{V}{g}\frac{dV_{Doppler}(r)}{dr} \qquad r\in[R_{min},R_{max}] \qquad (1)$$

where $V_{Doppler}$ is the velocity of the precipitation particles or aerosols at range r, measured relative to the aircraft, and equals the difference between the inertial speed vector $\vec{V}$ and the wind speed vector at range r $\vec{V}_W(r)$, projected

onto the radar beam as follows:

$$V_{Doppler}(r) = (\vec{V} - \vec{V}_W(r)) \cdot \vec{e} \qquad (2)$$

The vector $\vec{e}$ is the radar beam direction vector. Once a certain threshold $F_{crit}$ has been exceeded the area of $F>F_{crit}$ will be displayed on the weather radar display using a particular type of marking, see Figure 3. The basic value of $F_{crit}$ was set initially at -0.13 as required by the FAA[16]. Simultaneously with this type of windshear detector, also a reactive sensor may be present or not, depending upon the experimental set-up. Its $F_{crit}$ value was set at -0.105 [14]. One of the experimental goals was to determine the added benefit of having a reactive sensor as well, and to determine possible confusing operational conditions arising from these two different types of sensors.
Associated with both the laser and the ADR icon display is a caution and a warning area relative to the aircraft, see Figure 4a for the laser and Figure 4b for the ADR. These areas will be used to generate the appropriate alerts, depending upon the position of the icon(s).

Contrary to the drawback of the laser, the Doppler radar windshear detection function works fine in precipitation, but its windshear detection function will be lost in case of a "dry" microburst. New airborne radars, however, are able to still detect precipitation levels down to -10 dB, and this has also been assumed to be the case in this research investigation. This detection threshold of -10 dB amounts to a very small amount of precipitation of about 0.01 mm/h.

## 5 Go-around procedures evaluated

Any new sensor introduces new questions, such as how to cope with the information, and what to do and, if so, how to do it more optimally. Secondary goals in the 2 experiments described here therefore were to evaluate flight procedures. In the 1994 experiment, in case of no icon display, especially for the approach part a particular type of speed increment procedure was tested in conjunction with alerts coming from the windshear detection systems (i.e. laser and/or reactive sensor together). Concerning the go-around part, in case of a windshear warning a flight-director, based on the WTA type of go-around, was compared with a standard one in order to see if the WTA type of (constant) pitch steering, in conjunction with no configuration changes (i.e. a *fixed* aircraft configuration), would provide better flight safety during the go-around.
In the 1996 experiment a comparison was made between the *fixed* configuration (WTA type) and *varying* configuration (i.e. retracting flaps and landing gear on schedule) in the go-around. By having a full test matrix of such configuration changes across sensor types (i.e. windshear radar (ADR), normal weather radar with a reactive sensor, and both) an evaluation could be made of the flight safety of the best mix.

## 6 Alerting aspects

A difficult area concerns the integration of the various sensor alerts. Both visual cues (i.e. labels on the primary flight display) as well as aural cues (voice alerts) were used to alert the crew of an approaching windshear hazard.

Both visual and aural alerts were generated in accordance with the requirements of the FAA[14,15,16], JAA[9] and the recommendations from the SAE [8]. Visual windshear alerts depended upon the type of windshear sensor which detected the windshear (i.e. *sensor-driven* concept), and consisted either of a label "WINDSHEAR AHEAD" presented on the *lower* part of the PFD (forward-looking sensor, e.g. laser or ADR) or the label "WINDSHEAR" presented on the *upper* part of the PFD (Reactive sensor). In case of an *advisory* alert (only in the 1994 experiment) the labels were colored blue, in case of a *caution* alert the labels were colored amber, and in case of a *warning* alert they were colored red. Furthermore, depending upon the hazard level (see below), the alert labels were accompanied by identical aural cockpit alerts (e.g. "*windshear ahead*"), preceded by either the words "*advise*", "*caution*" or "*warning*", in case of an advisory, caution or warning alert respectively.

### 6.1 1994 experiment

The hazard levels used in the alerting concept used in the experiment of 1994 deviated from the internationally accepted Flight deck Alerting System (FAS) [8], and are given in Table 1.
The deviation is in hazard levels 2 and 3, where the caution alert used is based on negative performance rather than the performance-increasing trend used in the FAS. This was done in order to bring the piloted simulation results in line with numerical simulations done by the Flight Mechanics Action Group FM(AG05) of GARTEUR[10]. This actually brought the hazard level 2 closer to FAS level 3, as with a negative performance there is less time to act than in case of positive performance, and

**TABLE 1** *hazard levels compared to standard FAS*

| hazard level | FAS level | Alert type | Remark |
|---|---|---|---|
| 0 | 0 | Informa-tive | not used |
| 1 | 1 | Advisory | based on negative performance |
| 2 | 2 | Caution | based on negative performance |
| 3 | 2 | Caution | only used with a display |
| 4 | 3 | Warning | |

the meaning "caution" (i.e. no immediate action required) could no longer be valid. A new definition is hazard level 3, made similar to FAS level 2. In fact a FAS level 3 warning situation is present, however, the hazardous icon shown on the display is still outside the warning area (see Fig.4a), and then only a caution alert is given to the crew. The particular icon is also
colored amber/red (see Fig.1). The moment the FAS-level 3 icon penetrates the warning area a warning alert (i.e. hazard level 4: "immediate action") will be given. The icon color at this time also becomes solid red.

The aural cockpit alerting concept evaluated was *threat-level* based. Even though a sensor alert might disappear, the hazard level status would not be reduced during the simulated flight. That is, once a certain hazard level has been set by a sensor (e.g. 1: advisory) then only in case of higher hazard level alerts (e.g. warning) a voice warning will be passed on to the audio system. PFD alert labels were shown immediately as long as any one sensor detected a particular hazard, see also Figure 2.

### 6.2 1996 experiment

As far as the 1996 experiment is concerned the Flight deck Alerting System [8] approach was followed for the reactive sensor. For the ADR sensor the hazard levels 2 and 4, as defined in Table 1, were used. In case of a combination of the ADR and the reactive sensor, the reactive sensor's caution alert (based on positive performance) was suppressed, as is common practice nowadays. For the PFD alert labeling the same threat-level concept was used as for the 1994 experiment.

## 7 Risk calculation model

### 7.1 Risk assessment

An important tool for evaluating the flight safety of a particular configuration was the use of the NLR-developed risk assessment model, first used in the experiment in 1993[1], [2]. Compared to 1993 the model was adapted slightly in order to take into account the flight risk in the go-around arising from possible contact with the ground. Based on calculating the mean and standard deviation of several flight parameters, use of altitude above terrain as such would lead to erroneous results for the go-around phase, as in case of a normal climb-out the more or less linear trend in altitude would lead to a large standard deviation, and hence a large calculated risk of hitting the ground*. In the upgraded model this was circumvented by using the *mean* climb rate, in combination with the standard *error*. This turned out to work acceptably well: for all cases where a crash occurred the calculated flight risk also equaled 1.0 or very close to it. Furthermore, in the 1996 experiment, the risk model was further enhanced by also including the landing roll-out phase, calculating the risk of running off the runway for example.

### 7.2 General risk model

The general risk model used is comprised of the risk on the approach, the risk during the landing, or the risk in the go-around segment (the go-around and landing events are mutually exclusive), to calculate the total risk $R$ as follows:

$$R = R(Ap \cup Lnd \mid \overline{GA}).P(\overline{GA}) + R(Ap \cup GA \mid GA).P(GA) \quad (3)$$

The first term denotes the risk during approach or landing, given that no go-around has been made, and the second term denotes the risk during approach or go-around, given that a go-around has been made. In case a go-around has been made we have the probability:

$$P(GA) = 1 - P(\overline{GA}) \equiv 1 \quad (4)$$

According to the Bayesian rule we also have in general for the union of sets in Eq.(3):

$$\begin{aligned} R(Ap \cup Lnd \mid \overline{GA}) &= R(Ap \mid \overline{GA}) + R(Lnd) - R(Ap \cap Lnd \mid \overline{GA}) \\ R(Ap \cup GA \mid GA) &= R(Ap \mid GA) + R(GA) - R(Ap \cap GA \mid GA) \end{aligned} \quad (5)$$

In case the risk of an event on the approach, '$R(Ap)$' is independent of the risk of an event at landing, '$R(Lnd)$' or go-around, '$R(GA)$' (as is assumed here), then one can write for the right-hand terms in Eq.(5):

$$\begin{aligned} R(Ap \cap Lnd \mid \overline{GA}) &= R(Ap \mid \overline{GA}).R(Lnd) \\ R(Ap \cap GA \mid GA) &= R(Ap \mid GA).R(GA) \end{aligned} \quad (6)$$

The risk probabilities can be further defined for the approach, landing and/or go-around phases [1], [2].

The landing risk '$R(Lnd)$' was calculated differently for the experiment in 1996 than for the one in 1994. In the 1994 experiment only touchdown *event* data were used in the probability calculations (i.e. one event per landing), whereas in the 1996 experiment the landing roll-out segment data *statistics* were used to calculate the probabilities to exceed various parameters, using normal distribution functions [1], [2].

### 7.3 Risk conditions

The above risk probabilities were calculated, based on a number of conditions which were stipulated to be a risk. For the airborne part of flight the safety risk contained the risk to exceed the following terms:

- stall margin $\alpha_{mrgn} < 0$
- max. airspeed Va > 190 kts (flaps 30 deg; approach segment) or 238 kts (flaps 10 deg; go-around segment)
- load factor Az>+2.5g or < 0g (structural limits)
- glide slope deviation<-2 dots (on the approach only), or
- *mean* climb rate <0 (only the go-around segment).

In case of touchdown (TD) or landing/roll-out, the

* Use is made of the normal distribution to calculate risk probabilities

following conditions were defined to be a risk when exceeded:

- landing before touchdown point, i.e. $x_{TD} < -243$ m
- risk of runway overrun at the end. i.e. $x_{max} > 2475$ m
- risk of landing/getting off the runway, $|y_{TD}| > 16$ m
- tail strike, $\theta_{TD} > 11$ deg
- engine pod strike with the ground: $|\phi_{TD}| > 8.5$ deg
- exceeding max. wheel speed, $GS_{TD} > 203$ kts
- exceeding vertical speed at ground impact, $w_e > 600$ ft/min (landing gear structural limits).

The risk of exceeding the vertical acceleration at ground impact had been deleted, as it correlated strongly with the vertical speed at touchdown, making the two probabilities dependent. In fact, using one criterion will suffice. Note that some of the above values depend also on the aircraft type being used in the simulation experiment.

## 8 Conduct of experiments

This chapter will describe in more detail the conduct of the experiments, the experimental design and test matrix. In the 1994 experiment, performed in collaboration with the GARTEUR Flight Mechanics Action Group FM(AG07) members of ONERA, DLR and NLR, use was made of four national crews and two international crews (one from France and one from Germany). Four Dutch crews were used in the 1996 experiment. Both experiments concentrated on the approach and landing (or go-around) phase. Take-offs in windshear, although interesting enough, have not yet been considered due to limited resources. Due to the sheer size of the amount of data a limited selection of data was taken for further analysis and discussion in this paper.

### 8.1 Experimental design for the 1994 experiment

The experimental factors which play a role here are:

- *icon display* (yes, no). The combination of icon display *without* speed feedback was also called *display type A*.
- *speed feedback* (no, yes). The combination of speed feedback *with* the icon display was also called *display type B*. Also without the icon display the speed feedback mechanism was functional, in that it raised the alert thresholds, hence "delaying" the timing of the alerts.
- *rainfall* (no, yes). In case of rainfall the average intensity was about 30 mm/hr, and was concentrated within the vortex-rings of the microbursts. The rainfall rate of 30 mm/h was a reasonable amount of precipitation of "every-day life", yet would still be enough to influence the performance of the lidar such, that the effective look-ahead range, when fully immersed in this precipitation, would degrade from about 8000 m to about 2000 m. Rainfall was an experimental variable only in case an icon display was present, as without a display rainfall would not imply a "visual" awareness to the flight crews.

Concerning the objectives a sub-selection was made. In order to limit the scope of this paper the first two factors mentioned above were combined into the following factor and levels:

- *icon display type* (none, type A, type B).

The icon display type factor was used as a repeated measures, or within-subjects factor in a repeated measures ANalysis Of VAriance (ANOVA) test, to determine the significance of the effect. The repeats were made within crews.
In the work done by GARTEUR FM(AG07) additional experimental factors were evaluated [11], but these will not be discussed here for the sake of brevity.

A set of 4 weather scenarios was used, with the definition based on the location of the MicroBurst (MB) ***position*** relative to the approach path (*'no-threat'* [i.e. icon(s) outside the alerting zone] or a *'threat'* position [i.e. icon intersecting the alerting zone]), and the ***strength*** of the MB (*'severe'* or *'not severe'*), classified as given in Table 2. This factor of 'weather scenario' was used as a grouping factor in the ANOVA.

**TABLE 2** *Weather scenarios as function of MB position and strength*

| MB position | MB strength | weather scenario |
|---|---|---|
| no threat | not severe ($F_{av} \geq -0.10$) | ***benign*** |
| | severe ($F_{av} < -0.10$) | ***fly-by*** |
| threat | not severe ($F_{av} \geq -0.10$) | ***penetrate*** |
| | severe ($F_{av} < -0.10$) | ***worst-case*** |

Each class of windshear scenario contained several weather cases, each one having been made up of one or more microbursts at various locations and strengths, so that in fact there was one single classification but not one isolated weather case offered to the crews.

### 8.2 Experimental design for the 1996 experiment

In the 1996 experiment several windshear sensor options, in combination with 2 go-around procedures, were tested. The main experimental repeated measures factors of interest, and the levels, were:

- *windshear sensor type*: 1) reactive sensor with an ordinary weather radar (WX+R), 2) the Airborne Doppler Radar (ADR) only, and 3) the ADR ***with*** the reactive sensor combined (ADR+R).
- *type of go-around*: 1) WTA-type or 2) Standard type. In the first case the pilot was NOT allowed to change the aircraft's configuration but to keep it fixed while going around after a windshear alert, unless out of windshear and above 1000 ft. In the second case he had to retract flaps (to 20 degrees) as soon as possible as well as the landing gear as soon as the flaps were at 20 deg,

regardless of whether or not a windshear existed at the time of go-around. He was allowed to delay retracting flaps only in case of imminent ground impact or when below reference speed.

- *type of escape maneuver* (straight, turning). An effort was made to test the effect of making a turning escape maneuver in the go-around, rather than to go around straight, in the presence of shears. This factor turned out to be difficult to control, as in some cases where crews were supposed to go straight they turned, in defiance of "ATC instructions", while in other cases, where they were expected to make a turn in the go-around, they did not (apparently they got used too much to not having been allowed to make turning escape maneuvers that they didn't ask for it). As result of this only 4 valid cases (crew and weather combinations) remained, within which a full repeat was made, with in one case a straight go-around made, and in the repeat case a turning go-around. One case was flown with the ADR sensor, the other 3 cases were flown with the ADR+R sensor combination.

As grouping factor was used:

- *rainfall intensity*: dry (< 20 dBZ), medium wet (20 - 40 dBZ) and (very) wet (> 40 dBZ). Four of the FAA weather cases had rain core reflectivities which were sometimes above 50 dBZ(!) Of special interest of course were those cases where there was a dry microburst, to see what advantage the combination of the ADR and the reactive sensor might have.

Other factors not discussed here were the effect of rainfall on the deterioration of the airplane's aerodynamics, and another labeling concept to be shown on the Primary Flight Display (PFD).

Although FAA weather cases were used exclusively in the experiment of 1996, also here a weather classification analogous to the experiment of 1994 was performed, using the factors '*FBAR* * *position*' and '$FBAR_{max}$ *level*', as given in Table 3.

The factor '*FBAR pos*' determines the position of the center of the FBAR>0.13 contour which may show up on the radar. A threat was defined as an icon infringing the approach path. No-threat icons were positioned using subjective judgement. The '$FBAR_{max}$ *level*' determines the overall maximum FBAR that may occur, i.e. whether FBAR is above or below the maximum critical hazard level 0.13, above which an icon may be present on the ADR radar. Using this classification table allows these wind scenarios to be coupled to the wind scenarios of the second experiment.

* FBAR is the 1km averaged radar-F from Eq.(1).

**TABLE 3** *Weather classification as function of FBAR position and max. level*

| *FBAR pos* | $FBAR_{max}$ *level* | weather classification |
|---|---|---|
| no-threat | no hazard ($FBAR_{max}$<0.13) | ***benign*** |
| | hazard ($FBAR_{max}$≥0.13) | ***fly-by*** |
| threat | no hazard ($FBAR_{max}$<0.13) | ***penetrate*** |
| | hazard ($FBAR_{max}$≥0.13) | ***worst-case*** |

## 9 Data collection

The advantage of having a simulator is the easy access to a host of "flight data". The various performance data were processed into statistical measures such as mean, maximum, minimum and standard deviation of the performance parameters for several flight segments. The flight segments of interest in this paper were the following:

| segment | altitude range |
|---|---|
| approach | from 1800 ft to 50 ft or moment of go-around |
| go-around | from moment of go-around initiation to moment of altitude capture |
| landing | from moment of touchdown event (exp.2), to end of roll-out (exp.3) |

In this paper the objective results will be given in three dimensions, viz. "*performance*", "*safety risk*" and "*work-load*". From the many variables available for "*performance*" the choice made here was to use the *min. stall margin* during flight (i.e. on the approach or go-around if made), which correlated well with airspeed for example. Safety risk was calculated using the risk model (§7), and workload was determined from the pilot/copilot questionnaires using the McDonnell rating scales[17] to judge the demand on the pilot.

Subjective results are related to such things as situational awareness, usefulness of the (icon) display, acceptance of procedures, etc. All these data were collected from the questionnaires and were processed statistically.

The statistical significance (p-level) of a particular test or result represents the probability of error that is involved in accepting the observed results as valid, i.e. as "representative of the population". The p-level of 0.01 for example indicates a 1 percent probability (1 in 100 cases) that the relation between the variables found in the sample is based on chance. The significance levels for the various statistical tests used were p=0.01 (1 in 100) in case of highly significant differences, p=0.05 (1 in 20) in case of a significant difference, and p=0.10 (1 in 10) in case of a

weakly significant difference.

In order to evaluate the human performance, workload measures were given by both the Pilot-Flying (PF) and the Pilot-Not-Flying (PNF), so as to allow the crew's workload to be evaluated, i.e. PF and PNF together.

Also other questions were asked about the nature of the alerts, type of displays, etc. using quite extensive questionnaires. Important variables measured or determined were the go-around rate, type of alerts and timing of alerts, and other crew-related functionalities.

## 10 Results and discussion

The results of the experiments will be discussed regarding the three dimensions for each of the main factors of interest mentioned before.
Furthermore, in case go-arounds were made only straight go-arounds were selected, so as to avoid having additional variability in the performance due to having made turning escape maneuvers or not. Also only "dry" runs were taken, except when rainfall was an experimental factor.

### 10.1 Laser-driven icon display and rainfall

From the wealth of information and data available, a sub-selection was made to portray the results concerning the laser-driven icon display *type* (i.e. icon display and speed feedback with the display), together with the results of the experimental factor of rainfall intensity.

10.1.1 Effect of icon display type. The main effect of the icon display type on the ***min. stall margin*** turned out to be quite significant, in a statistical sense ($F(2,30)=14.10$; $p<.0000$). The mean values of the min. stall margin as well as the go-around rate are given as function of the type of icon display (none, type A or type B) in Figure 5. As the figure shows especially the type A display had the highest stall margin and lowest go-around rate, and in case of no display the min. stall margin was worst, and the go-around rate was the highest. Breaking this down per weather scenario, see Figure 6, shows that the improvement in stall margin due to the display occurred with all scenarios except the penetrate scenario, and the best improvement occurs with the benign weather scenario. Across the board display type A looks like the "best" display type, with an exception perhaps for the worst-case scenario, where display type B (i.e. *with* speed feedback) had a slightly better (min.) stall margin.

The effect of the icon display type on ***safety risk***, per weather scenario, is shown in Figure 7. Although overall the display type effect was not significant, it was significant ($F(2,30)=5.339$; $p<.0104$) for the "fly-by" weather scenario. As Figure 7 shows the worst risk for this scenario was obtained in case of no display, while display type A again had the best, i.e. lowest safety risk. For all the other weather scenarios there was no significant difference due to icon display type. Apparently it was just for this "fly-by" scenario where the presence of the icon display made a big difference in circumnavigating windshear. As the go-around rate shows (see later) the improved safety risk was obtained at the same time with a zero go-around rate. The worse risk for the no-display case apparently was caused by going around instead of landing!

The effect of the icon display type on the ***crew's workload*** is given in Figure 8 as function of display type, and in Figure 9 as function of weather scenario, per display type. The type of display had a significant main effect on the crew's workload ($F(2,6)=3.29$; $p<.0438$), see Figure 8, where the workload for display type B (i.e. icon display *with* speed feedback) was higher than for display type A, or even no display. As the speed feedback tended to drive crews to continue the approach, and to go around much later than in case of display type A, it was generally harder work for the crews to negotiate the shear. Figure 9 makes it evident that much was to be gained by adding an icon display (of type A) especially for the benign weather case. Display type B as well as no display yielded the same level of workload. Apparently the speed feedback mechanism drove the pilot's behavior closer to that for the no display situation (i.e. to continue flight towards a possible threat), and as result did not improve the workload. Addition of an icon display was also accompanied by a great reduction in go-around rate (see Figure 11 later). For the worst-case scenario, however, the type A icon display yielded a much higher workload than for type B or no display at all. In all cases go-arounds were made anyhow, but in case of display type A the crews also saw the ominously looking icon staring in their face, whereas with type B the threat may have looked less ominous due to changes in the color of the icons.

An interesting variable to compare workload with is the ***in-situ* maximum windshear hazard** (min. $F_{av}$) which the airplane experienced during flight. This variable is shown as function of weather scenario, per display type, in Figure 10. This parameter was determined by the on-board reactive sensor. Please note that a negative value of $F_{av}$ means performance loss, and vice-versa. Overall the type of display had a highly significant main effect ($F(2,30)=8.13$; $p<.0015$) on min. $F_{av}$, with the lowest (average) max. hazard ($F_{av}=-.05$) for display type A, and the highest hazard ($F_{av} = -.20$) in case of no display, and type B in between ($F_{av} =-0.13$). Figure 10 shows that the greatest hazards occurred with the worst-case scenario without a display (obviously), but also with the benign scenario(!). Looking back, Figure 9 reveals that the workload is proportional to the max. windshear hazard in case of the benign weather scenario, but is *inversely* proportional to the max. windshear hazard in case of the worst-case scenario. Apparently the higher workload for display type A in the worst-case scenario led to a better, i.e. lower windshear hazard, while in case of the benign scenario it was mainly the visual information of the weather which led pilots to believe it was an easy, low-workload situation, however with some surprises! Notice

though, looking at Figure 7, that none of this correlated with the safety risk, however, which showed worst risk for the no-display situation in case of the fly-by scenario instead.

Another interesting variable to judge the effect of the display type is the ***go-around rate***, shown in Figure 11 as function of weather scenario, per display type. For the more "hazardous" scenarios, like the worst-case and penetrate scenarios, there is not much difference due to display, but especially for the more "complex" scenarios, in terms of decision making, there is quite an effect of display. In case of the benign scenario the go-around rate dropped from 100 percent to zero for the type A icon display, a great improvement when considering the safety risk (Fig.7), stall margin (Fig.6) or operational implications. In case of the fly-by scenario both display types (A and B) reduced the go-around rate (from 80 percent to zero), as intended since it was supposed to be a "fly-by" situation. The reduction in go-around rate was even accompanied by a (slight) increase in the min. stall margin (Fig.6), and a significant reduction in safety risk (Fig.7). In the case of the penetrate scenario, however, one can notice an increase in go-around rate from zero to about 20 percent, when adding an icon display. Considering the amount of runs (6), this amounted to one go-around for type A and one for type B. From an operational point of view this could be interpreted as two go-arounds too many because of the addition of an icon display. Closer analysis indicated the go-arounds were made because of the aircraft becoming unstable on the approach, and not because of displayed windshear features.

In this comparison of icon display types within crews it turned out that the pilot's rating of the ***usefulness of the icon display*** type did not differ statistically significantly between display types, although the usefulness of icon display type B was rated slightly better than that of type A.

When ranking the various display types among one another in terms of ***pilot's preference***, using Saaty's method [12], then having a display was preferred far above not having a display, and display type A ranked slightly better than type B, see Figure 12, instead of less, as indicated above. This latter was explained by the pilots in that, in hindsight, they did not like the speed feedback with the display, as it drove them towards continuing the flight into a weather situation which normally would have compelled them to go around. Making a go around in such a once-in-a-lifetime situation was considered quite acceptable to them.

10.1.2 Effect of rainfall. The effect of rainfall was initially observed to be not as dramatic as it turned out to be. A more detailed analysis indicated that rainfall had a much greater effect on the efficacy of the icon display than hitherto assumed. In Figure 13 the effect of rainfall on the ***min. stall margin*** is shown as function of weather scenario. Rainfall had a significant main effect ($F(1,10)=6.70$; $p<.0270$). On average the min. stall margin became one degree less with rainfall, and Figure 13 shows that this occurred especially with the penetrate and worst-case scenarios, where in the latter case the min. stall margin reduced significantly ($p<.0686$) from 7 deg without rainfall, to 6 deg with rainfall. Although the change in min. stall margin was even larger for the penetrate scenario, the difference turned out not to be significant (too much spread in the data).The go-around rate for this latter scenario also went up, from about 20 percent to 100 percent (see later). Notice though, that for the benign weather scenario the min. stall margin improved when having rainfall, viz. from 8 to 11 degrees, while the go-around rate dropped somewhat.
There was also a significant ($F(1,9)=3.92$; $p<.0792$) interaction between display type (A, B) and rainfall, see Figure 14. The performance with icon display type A showed a much greater sensitivity to rainfall than it did with type B. The min. stall margin dropped from 9.5 deg to 7.5 deg due to rainfall for type A, but remained unchanged at about 8.5 deg for type B. Apparently the "filtering" effect of the speed feedback concept in the type B icon display made it also less sensitive to the effect of rainfall, which would have a somewhat similar effect on visual appearance as speed feedback, viz. a reduction in icon size, sometimes with icons disappearing completely because of a reduction in the detection range.

Also ***the go-around rate*** for the same conditions as for the min. stall margin is shown in Figure 15. It becomes clear that the reduction in min. stall margin for the penetrate scenario is accompanied by an increase in go-around rate, from zero to 100 percent. Partly the effect of rainfall in terms of min. stall margin and go-around rate can be attributed to reduced visual cues (rainfall also reduced the visual slant range), leading crews to go around more often, but also to a reduction in size of the icons, inducing pilots to believe the situation was much more benign than it really was. As this was a "penetrate" type of scenario anyway, the effect of rainfall also led to a greater degree of penetration of shears which, in combination with the reduced visual cues, would lead to a greater go-around rate at a later moment in flight.

It turned out that rainfall had no significant effect on ***flight safety risk*** whatsoever. In one case the flight safety risk value reached a value of 0.96 (worst-case scenario, in rainfall, with type A display), which was due to the mean vertical speed in the go-around becoming just negative. It turned out that in this run the max. angle of attack exceeded stall, i.e. the min. stall margin was -1.1 deg, but no crash resulted from this. When applying a non-parametric test, in order to cope with this outlier, then rainfall was no longer a significant factor. When deleting this case, then rainfall also was no longer a significant factor. More data is needed to be able to be more specific.

Rainfall did have a significant main effect ($F(1,22)=7.56$; $p<.0117$) on the ***crew's workload***, see Figure 16.

Especially for the penetrate scenario the workload increased with rainfall to even above the level for the worst-case scenario, whereas for the benign scenario it reduced with rainfall, although only slightly. It is obvious that penetrating shears with rain is a higher workload situation than without rainfall, as especially the visual cues "at the end" are important for judging the aircraft's position relative to the runway, in order to decide whether or not to go around. As Figure 15 shows the rainfall condition in general resulted in a higher rate of go arounds, especially for the penetrate scenario.

Pilots did not notice a reduced usefulness of the icon display due to rainfall.

Based on the performance, safety risk, crew workload and pilot's preference the best type of icon display is type A, i.e. the laser-driven icon display WITHOUT speed feedback.

10.2 Radar-driven icon display

In the case of the 1996 experiment it was especially the windshear radar icon display (WX+Reactive sensor combination as the baseline, the ADR and the ADR+Reactive sensor) and the type of go-around (*WTA* or *Standard.*), which were the major parameters of interest. The runs were flown on a number of weather scenarios, which included rainfall, expressed in terms of the max. dBZ level that existed in this scenario. In case of a go-around only straight go-arounds were selected.

10.2.1 Effect of sensor type. The effect of the type of sensor on performance (i.e. ***min. stall margin***) is shown per weather scenario in Figure 17. The type of sensor had a significant main effect on the min. stall margin ($F(2,12)=6.15$; $p<.0145$), with the highest min. margin for the ADR sensor (average 8.8 deg), and the lowest min. margin for the WX+R sensor combination (av. value 7.6 deg). It is clear that for the two "easy" scenarios (the benign and the fly-by scenarios) there is no difference between the sensors, but especially for the more difficult ones there is a difference. For the penetrate scenario the ADR sensor yielded an outlying better ($p<.0555$) min. stall margin than for the other two sensor types. This was achieved at the same time with a lower go-around rate (see Fig.18). So here continuing the approach apparently led to better stall margins than by aborting it. For the worst-case scenario both ADR type of sensors performed better, at the $p=.000112$ level, than the baseline WX+R sensor combination. This improvement turned out to be even more pronounced for the *Standard* type of go-around (about 3 deg) than for the *WTA* type of go-around (about 1.5 deg).

The effect of the sensors on the ***go-around rate*** is shown in Figure 18. As expected the go-around rate is very low, and even zero for the ADR sensor, for the "easy" scenarios. For the penetrate scenario with ADR sensor the go-around rate reached a low of about 20 percent, whereas for the other two sensors the go-around rate went up to about 80 percent. For the worst-case scenario the go-around rate was 100 percent, as expected. Despite the same go-around rate for this scenario, a better stall margin for the ADR-type of sensors (Fig.17) could only be the result of having gone around earlier, as was confirmed by checking on the moment of go-around. The average go-around distance from touchdown for this scenario, based on 3 crews, was about -3000 m (i.e. *before* touchdown) for the WX+R sensor, but for the ADR sensor it was -5500 m, and for the ADR+R sensor combination it was -5000 m (i.e. a little later than with the ADR). This difference in go-around distance between the WX+R and the ADR-type of sensors of 2-2.5 km accounted for the worse min. stall margin for the WX+R sensor combination. The closer go-around distance of -3000 m for the WX+Reactive sensor was due in part to dry microbursts, which were undetectable by the WX radar alone. The reactive sensor gave a later warning in this case than the ADR sensor did.

The type of sensor did not have any significant main effect on ***flight safety risk*** whatsoever, however, for one particular weather scenario it did. In Figure 19 the flight safety risk is shown as function of weather scenario, per sensor type. For the worst-case scenario the flight safety risk for the WX+R sensor combination ($R=4.0*10^{-4}$) was significantly higher ($F(1,6)=4.656$; $p<.0743$) than for the ADR or the ADR+R sensor ($.01-.40*10^{-4}$). Also for the benign scenario the WX+R sensor combination scored much worse than for the other sensor types, however this time the difference was (just) not statistically significant ($F(1,6)=3.368$; $p<.116$) due to more spread. More data is needed to be conclusive about this.

The ***in-situ*** **max. windshear hazard** (min. $F_{av}$), as experienced by the aircraft, is shown in Figure 20. Clearly one can see the benefit of the ADR icon display types by the drop in min. $F_{av}$ for the worst-case scenario from a value of -0.20 down to below the alert level (-0.13). This alert level would trigger the ADR alert logic. Notice that for the worst-case scenario the max. windshear hazard, as experienced (on average) by the airplane in case of an ADR type of sensor, never was above the warning alert level of -0.13, so the warning alert from the ADR, which triggered the crews into making a go-around, was a timely one.

The effect of the type of sensor on the ***crew's workload*** is shown in Figure 21 (left-hand axis), together with the go-around rate along the right-hand axis. The type of sensor had a significant main effect ($F(2,24)=4.03$; $p<.0310$) on the crew's workload. It was especially the workload with the ADR sensor which was significantly ($p<.0746$) more than for the other sensors, accompanied by a lower go-around rate. A closer inspection revealed that it was especially for the worst-case scenario where the workload with the ADR sensor was significantly higher ($F(1,12)=5.659$; $p<.0348$) than for any other sensor or sensor combination. Also, especially for the benign scenario, the workload for the ADR+R sensor was

significantly lower ($F(1,12)=6.088$; $p<.0296$) than for the WX+R sensor.

Furthermore there was a highly significant ***interaction*** between windshear sensor and the type of go-around, $F(2,24)=7.32$; $p<.0033$, see Figure 22. This was because in the specific case of the ADR+R sensor, the workload for the *WTA* type of go-around was significantly lower ($F(1,12)=28.49$; $p<.000177$) than the workload for the *Standard* type. Further probing revealed that this was especially due to the Pilot-Not-Flying (PNF). This can be explained by the fact that for this type of sensor the average go-around distance was about 700 m less (see §10.2.1) than for the ADR sensor alone. In this case, carrying out a *WTA* type of go-around meant for the *co*pilot to not retract the flaps and landing gear, as opposed to the *Standard* type of go-around where he had to perform these duties, with a resulting higher workload. Retracting flaps and the landing gear in such a "high stress" situation (i.e. a fairly late go-around) apparently meant a much higher *mental* load for the PNF. The workload of the PF remained fairly constant, as the mental workload of his flying task did not change much.

Also there was a highly significant effect (Kruskal-Wallis test $H(2,N=96)=9.017$;$p=.0110$) on the ***situational awareness*** for the fly-by and the worst-case weather scenarios. The situational awareness is shown in Figure 23. For these two scenarios, which are the most "interesting" ones from a situational awareness point of view, the situational awareness scored "fair-to-good" for the WX+R sensor, which increased to "good" for the ADR and the ADR+R sensor combination. Obviously the presence of windshear icons, through the ADR feature, helped increase the situational awareness.

Also, for the same weather conditions as above, the non-parametric Kruskal-Wallis test showed the ***weather radar's display usefulness*** to be better for the ADR and ADR+R sensor ("fair") than for the WX+R sensor ("bad-to-fair") ($H(2, N=96)=9.786$; $p=.0075$). Mean values of the usefulness of the weather radar display are given in Figure 24. It is, for the same reasons as above, obvious why the usefulness was rated better for the ADR type of sensor.

10.2.2 Effect of type of go-around. The effect of the type of go-around on the ***min. stall margin*** is shown in Figure 25 as function of weather scenario. The highly significant main effect of go-around type ($F(1,6)=34.24$; $p<.0011$) was an average one degree ***lower*** min. stall margin for the *Standard* type of go-around, and this figure shows that this occurred especially with the penetrate (2.5 deg) and the worst-case (one deg) scenarios. It was not expected that this trend would emerge. The interaction between sensors and go-around type, see Figure 26, nearly becoming significant, further showed that this lower stall margin occurred especially with the WX+R sensor, where the min. stall margin was highly significantly lower ($F(1,6)=23.68$; $p<.00280$) for the *WTA* type than for the *Standard* type of go-around. In view of the effect of sensors it becomes obvious that the combination of WX+R sensor and *Standard* type of go-around is not the best one. This seems to support the notion that, despite the presence of a weather radar, the performance in windshear is driven mainly by the functioning of the reactive sensor only. Hence the currently applied *WTA* type of go-around for this WX+R sensor combination seems to be suitable.
The best overall min. stall margin was obtained with the ADR sensor, regardless of the type of go-around or weather scenario.

It turned out that the type of go-around had no significant effect on ***flight safety risk***, although the mean values tended to double from $1.0*10^{-4}$ for the *WTA* type of go-around , to about $2.0*10^{-4}$ for the *Standard* type of go-around. When including the results of the fourth crew then the effect did become significant. However, the fourth crew repeated the runs of the first crew, but with a lower critical hazard level ($F_{crit}=0.10$) set for the ADR warning threshold, so in this case the experiment set-up becomes unbalanced. More data runs (i.e. crews) are required in order to be able to decisively determine the effect of the go-around type on flight safety risk.

The effect of the type of go-around on the ***demand on the crew*** is shown in Figure 27 as function of weather scenario. A significant interaction between weather scenario and go-around type ($F(3,12)=3.67$; $p<.0437$) indicated a dependency of workload on the type of go-around, further depending upon the type of weather scenario, such as is suggested by the figure. The only significant difference occurred for the penetrate scenario, where the workload for the *Standard* type of go-around was significantly higher ($F(1,12)=6.984$; $p<.0215$) than for the *WTA* type. This increase in the crew's workload for this scenario was entirely due to a higher (mental) workload for the PNF, as explained before.

Another significant interaction existed between the type of go-around and crew member ($F(1,12)=17.24$; $p<.0013$), shown in Figure 28, which indicates that the workload for the PF was significantly lower for the *Standard* type than for the *WTA* type of go-around, while that of the PNF was higher. The PNF had to perform more duties, and in quick succession, in case of the *Standard* type of go-around. right after go-around initiation, putting (apparently) a higher mental stress on the PNF, while the PF only had to perform the same type of operation, but without the long wait for aircraft "clean up", hence with a lower mental load. This interesting feature of interaction between crew members could only have been determined by having queried both pilots.

An indication of the aircrew's ***go-around procedure acceptance*** was obtained from questions in the questionnaire, where each PF rated the acceptance of the particular go-around procedure (in case of a go-around), on a scale of from -4 (fully rejected) to +4 (fully accepted),

with 0 as the neutral point. In general both go-around procedures (*WTA* or *Standard*) were rated better than neutral and were not significantly different from one another, with a slight preference for the *Standard* type. It was expected that the *Standard* go-around procedure would be more acceptable with the ADR or the ADR+Reactive sensor, and that the *WTA* type of go-around would be more acceptable with the WX+Reactive sensor, however, this interaction between sensors and go-around type was non-existent. Therefore no clear pilot's preference about the combination of type of sensors and go-around procedures exists.

Although the interaction between the go-around type and weather scenario was not significant ($F(1,1)=12.35$; $p<.1765$) the trend it shows, see Figure 29, was that the *Standard* go-around procedure was accepted better in case of the penetrate scenario, where late go-arounds generally were made, while in case of the worst-case scenario there was no clear preference for any particular go-around procedure. It should be reminded that there were only a few complete repeats within the 3 crews available, so more data are required to be more specific. When including e.g. the results of the 4th crew then the *Standard* type of go-around procedure tended to be accepted better than the *WTA* type, regardless of the type of weather scenario. Pilot comments supported this trend, stating that they would prefer to make the *Standard* type of go-around procedure (i.e. retract gear and flaps) when warned well before the real windshear threat. When going around in the windshear itself, e.g. due to a reactive sensor alert, then they didn't mind leaving the aircraft's configuration unchanged (i.e. *WTA* type) as they climbed away from the threat. They did object, however, to climbing away early, while having to keep the flaps and gear down until 1000 ft AGL, as they preferred to accelerate faster.

10.2.3 Effect of precipitation level. The windshear cases per weather scenario were also classified in terms of the max. precipitation level as 'dry' (precipitation <20 dBZ), 'medium' (wet) (20-40 dBZ) and (very) 'wet' (>40 dBZ). The factor of precipitation level was not really independent, as with each level of precipitation within one weather scenario, a different windshear (i.e. microburst) case applied. Therefore, looking at differences in precipitation might also mean looking at differences between microbursts. In the 1996 experiment it turned out that, within the (6) repeats per crew that were made, there were from one to two windshear cases per weather scenario for each class of precipitation level, as indicated in Table 4 (only straight go-arounds were selected).

The effect of sensors per precipitation level on performance will at least be interesting to see, in order to determine if the functioning of the windshear sensor will be affected by the amount of precipitation.

In terms of ***min. stall margin*** the effect of precipitation, per sensor, is shown in Figure 30. In case of dry weather the type of sensor did not matter for the min. stall margin. For the medium wet precipitation case there was a significant effect of sensor, with the WX+Reactive sensor being the worst, and the ADR being the best. For the (very) wet precipitation case the differences between sensors became even more, with again the WX+Reactive sensor being the worst, and the ADR being the best. Maximum (average) difference in min. stall margin amounted to 2 degrees in this case. Especially the *WTA* type of go-around in the (very) wet case gave rise to low stall margins. It is apparent that the use of the ADR, or ADR+R, type of sensor is a big help in improving the safety by increasing the minimum stall margin relative to the baseline case (i.e. WX+R sensor). Reasons for this are obvious: more situational awareness, especially of the more dangerous shears.

TABLE 4 *No. of weather cases per precipitation level per weather scenario*

| precipitation level | weather scenario | no. |
|---|---|---|
| Dry (<20 dBZ) | benign | 2 |
| | fly-by | 2 |
| | penetrate | 2 |
| | worst-case | 1 |
| Medium (wet) (20-40 dBZ) | benign | 2 |
| | fly-by | 2 |
| | penetrate | 1 |
| | worst-case | 1 |
| (Very) wet (>40 dBZ) | benign | 2 |
| | fly-by | 2 |
| | penetrate | 1 |
| | worst-case | 1 |

The effect of precipitation, coupled with windshear sensor, on ***flight safety risk*** is shown in Figure 31. In case of dry weather there was no effect of the type of sensor on the flight safety risk. In case of medium wet precipitation, the ADR+R sensor combination had a significantly lower ($p<.0598$) flight safety risk than the other sensor types. Although for the wet precipitation case it looks as if there is a large(r) outlying risk value of $6.5*10^{-4}$ for the WX+R sensor, mainly attributable to the *WTA* type of go-around, this was not statistically significant, however ($p<.183$), although it does indicate a trend: the WX+R sensor combination, especially in case of the *WTA* type of go-around, yielded a higher flight safety risk than with the ADR type of sensors.

The effect of precipitation, per sensor type, on the ***crew's workload*** is shown in Figure 32. Only in case of medium precipitation the workload for the ADR+R sensor combination was significantly higher ($F(1,12)=6.707$; $p<.0237$) than for the other two sensors. This "upward" dip in the trend with rainfall for this type of sensor was mainly due the PF's workload for the *WTA* type of go-around being higher than for the other two sensor types and go-around type. Apparently waiting with the "clean up" of the aircraft, flying in a "high drag" situation (and lower

speeds), was not liked by the PF in particular.

The ***usefulness of the WX radar display*** also depended upon the amount of precipitation, as well as sensors, see Figure 33. Especially for the dry situation the ADR or ADR+R sensor (obviously) scored better than the WX+R sensor (Kruskal-Wallis test H(1, N=43) = 3.525; p=.0605), as in this case the ADR could still show windshear icons, whereas the WX+R did not show any weather at all. For the medium and wet precipitation levels the difference in usefulness between sensors was no longer significant, in a statistical sense. Overall the usefulness of the radar display increased significantly with precipitation, using the non-parametric Kruskal-Wallis test (H(2,N=186)=27.51; p=.0000), and was best for the (very) wet precipitation level.

10.2.4 Effect of escape maneuver

Generally, when making turning maneuvers the stall margin is expected to *de*crease. The effect of escape maneuver turned out not to be significant for the ***min. stall margin***, however, contrary to expectations, the min. stall margin *in*creased by about 1.8 deg when making turning escape maneuvers. Also concerning the ***flight safety risk*** the maneuver effect was not significant, but also here the flight safety risk improved somewhat for turning escape maneuvers. This trend was similar to what was found in the 1994 experiment (not reported here) [11], and to what was found by Visser [18].

There was a significant effect of escape maneuver, however, on the ***crew's workload*** (F(1,6)=5.83; p<.0523). In Figure 34 the effect of the escape maneuver is shown as function of crew member (PF, PNF). For turning escape maneuvers the workload was higher than for straight go-arounds. Although not significant, Figure 34 does show that the increase in workload for turning escape maneuvers was stronger for the PNF than for the PF.

10.2.5 Windshear severity experienced

For each run the pilot filled out in the questionnaire a measure about the windshear severity he experienced "during the flight", on a scale from 'none', 'light', 'moderate', 'heavy', 'severe' to 'extreme'. It is interesting to see how his subjective rating varies with the application of different sensors and/or go-around procedures, and how it compares with the *in-situ* max. windshear hazard measured by the reactive sensor (min $F_{av}$). For a selected data set of straight go-around cases (obviously on penetrate and worst-case windshear scenarios) the mean values for the windshear *severity* (subjective) and *hazard* (objective) are given in Figure 35. The open symbols are the subjectively determined severity values, and the solid symbols are the objectively determined hazard values. Circled symbols are for the WTA go-around procedure, while the delta symbols are for the Standard go-around procedure. It is striking to see the good agreement between the subjective and objective data, certainly for the ADR and the ADR+R sensor. For the WX+R sensor the pilot rated the windshear experienced with the WTA procedure less severe than the measurements indicate. In his subjective opinion the Standard go-around procedure and the WX+R sensor combination led to a more hazardous windshear situation than when making a WTA type of go-around, although this difference did not reach statistical significance, and also the objective data did not show this trend. For the ADR sensor there was no difference in subjective rating, while for the ADR+R sensor combination he rated the windshear severity less in case of the Standard type of go-around than for the WTA type of go-around, but also here this difference was not (yet) significant, in a statistical sense. This latter subjective finding is substantiated by the objectively measured data (min. $F_{av}$). In case of the Standard type of go-around the type of windshear sensor did have a significant effect on the subjectively rated windshear severity (Friedman's ANOVA $\chi^2$ (N=8, df=2)=5.7; p<.0579), and nearly so on the max. *in-situ* windshear hazard (F(2,14)=2.58; p<.111), with the ADR type of sensors yielding better, i.e. less windshear hazard or severity.

It is possible to draw a scale of windshear severity experienced by the pilot, versus the max. *in-situ* windshear hazard experienced. This has been done for all the cases, with the exception of the crash cases (they gave rise to deviant results). The result is given in Figure 36. A linear relationship has also been drawn in the figure, showing the fairly linear behavior of the pilot in rating the windshear severity, compared to the *in-situ* max. windshear hazard. The $F_{crit}$ level of -0.13 clearly separates the windshear class of 'moderate', and less, from the windshear class of 'heavy' and above. The large std. deviation shows, however, that there was quite some spread in the data, particularly in the light-to-heavy category.

## 11 Concluding remarks

Overall the results show improved flight safety and good usefulness of the modern windshear hazard displays evaluated. The speed feedback concept, i.e. display type B, was not accepted by the pilots, as it tended to drive them to continue flight into adverse weather, with a commensurate increase in workload. Making a go-around instead was considered a good alternative to such unsafe conditions. The flight performance with the laser-driven icon display was rather adversely affected by precipitation. The stall margin reduced, and also the go-around rate went up with precipitation. The Airborne Doppler Radar (ADR) was liked very well, providing better situational awareness and being more useful than the weather radar only. It was quite a "natural" sensor to have, partly due to its integration with the weather radar's display. Pilots liked to have windshear cues confirmed by weather cues. Obviously, the less hazardous the cues that were visible the less useful the icon display was rated.

Because of the limited amount of data no firm statements can be made regarding the pilot's preference for a particular sensor/go-around type combination. However, in view of the flight safety risk and the minimum stall margin

one can conclude that the WX+R sensor should not be coupled with the *Standard* type of go-around, but rather with the *WTA* type of go-around. The performance and flight safety risk improved with the ADR type of sensor. More data is required to validate these findings. Making turning escape maneuvers did not significantly increase the stall margin or reduce the flight safety risk, but the workload increased. Also here more data is required to validate this finding.

In case of integration of multiple sensors, such as a windshear radar and the reactive sensor for example, a careful tuning must be established in order to prevent loop holes in the detection of hazardous windshear. The best sensor combination is the ADR with the reactive sensor. The reactive sensor also detects the "dry" weather cases which may be missed by the ADR sensor, while the ADR sensor detects the wet windshear cases much more timely than the reactive sensor does. From a first impression of the crash cases which occurred in the 1996 experiment, and pending further analysis, it may be recommendable to not suppress the reactive sensor caution alert when combined with the ADR, even though its meaning (i.e. performance increase) conflicts with that of the ADR.

## 12 Acknowledgment

The authors wish to acknowledge the grateful cooperation of the pilots from KLM, Martinair, Air France and Lufthansa, as well as from the University of Technology Delft, The Netherlands and the C.E.V., France for their contribution. Also much appreciation is extended to the FAA and NASA for providing the windshear database, and to colleagues from ONERA and DLR in their contribution to the work performed . This research was in major part conducted under contracts awarded by the Netherlands Agency for Aerospace Programmes (NIVR) and the Netherlands Department of Civil Aviation (RLD), and partly under NLR's basic research program.

## 13 References

1. Haverdings, H. & Rouwhorst, W.F.J.A. (1994): "*Initial results of a piloted simulator investigation of modern windshear detection systems*". ICAS paper 94-7.1.3, 19th ICAS congress, Anaheim, Calif., USA, September 1994. Also NLR TP 95519 U.
2. Rouwhorst, W.F.J.A. & Haverdings, H. (1996): "*Some results of piloted simulator investigations on windshear detection systems and icon display concepts*". ICAS paper 96-3.8.3, 20th ICAS congress, Sorrento, Italy, Sept.1996. Also NLR TP 96592 U.
3. FAA (1987): "*Windshear Training Aid*". Example Windshear training Program, February 1987.
4. Douglas, P., Lewis, M.S., Hinton, D. (1996): "*Airborne systems technology application to the windshear threat*". ICAS paper 96-5.7.1, 20th ICAS congress, Sorrento, Napoli, Italy, September 1996.
5. Schultz, T.A. (1990): "*Multiple vortex-ring model of the DFW microburst*", Journal of Aircraft, Vol.27, No.2, February 1990.
6. Elmore, K.L. & Sand, W.R. (1989): "*A cursory study of F-factor applied to Doppler radar*", paper 3d Int'l Conf. On Aviation Weather Systems, Jan-Feb. 1989.
7. Haverdings, H. (1992): "*Functional description and software specification of an airborne laser windshear detector*". NLR Contract Report CR 90292 L, 1990.
8. "*Flight Deck Alerting System (FAS)*", SAE, Aerospace Recommended Practices ARP4102/4, Issued July 1988.
9. JAR 25, Amendment 93/1, 8 March, 1993.
10. Huynh, H.T, (ONERA) Descatoire, F. (ONERA), Hahn, K.-U. (DLR), König, R (DLR), Haverdings, H. (NLR), Rouwhorst, W.F.J.A. (NLR) (1996): "*(GARTEUR open) Performance evaluation of airborne reactive and forward-looking windshear detection systems on a simulated aircraft. Final report on Task 2 of FM-AG(05)*". GARTEUR TP-092, December 1996.
11. Haverdings, H. (NLR), Rouwhorst, W.F.J.A. (NLR), Hahn, K.-U. (DLR), König, R. (DLR), Descatoire, F. (ONERA), Huynh, H.T. (ONERA) (1997): "*(GARTEUR OPEN) Piloted investigation of windshear guidance, detection systems and hazard display concepts. Final report of FM-AG(07)*". GARTEUR TP-105, December 1997.
12. Saaty, T.L. (1980): "*The analytical hierarchy process*". McGraw-Hill, !980.
13. Switzer, G.F., Proctor, F.H., Hinton, D.A., Aanstoos, J.V.(1993): "*Windshear database for forward-looking system certification*". NASA Technical Memorandum 109012, November 1993.
14. Anon. FAA Technical Standard Order: "*Airborne Windshear warning and escape guidance system for transport Airplanes*", FAA-TSO-C117, 24 July 1990.
15. Anon. FAA Advisory Circular "*Criteria for operational approval of airborne windshear alerting*", FAA-AC-120-41, 11 July 1983.
16. Anon., FAA (1995) "*Airborne short and long range windshear predictive systems (Forward-looking windshear systems). Interim certification requirements*". Revision 10.2, January 1995.
17. McDonnell, J.D. (1968): "*Pilot rating techniques for the estimation and evaluation of handling qualitites*". AFFDL-TR-68-76, December 1968.
18. Visser, H.G. (1994): "*Optimal lateral escape maneuvers for microburst encounters during final approach*". Journal of Guidance, Control and Dynamics, Vol.17, No.6, Nov-Dec. 1994.

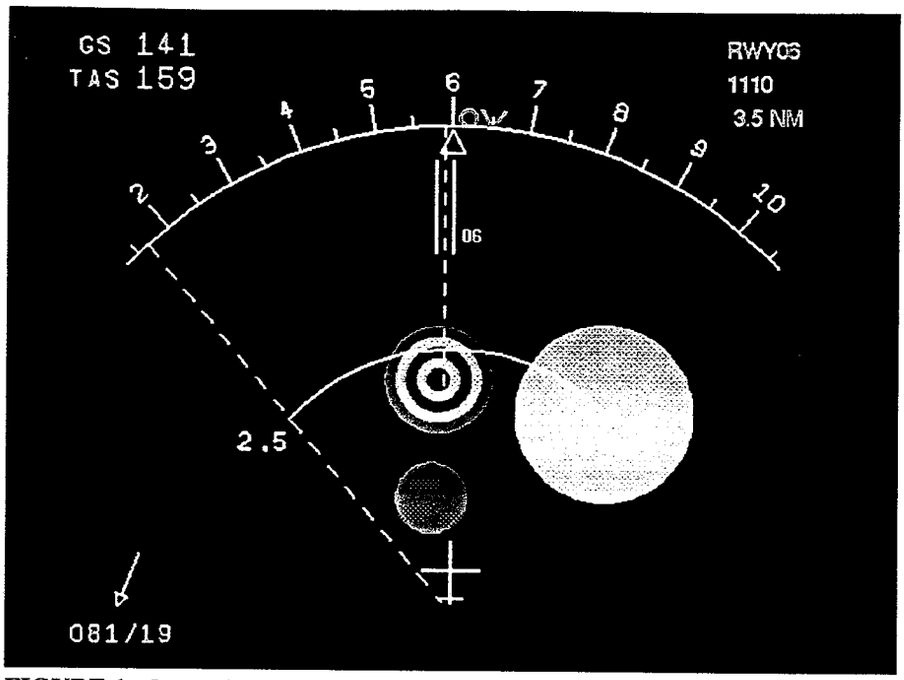


**FIGURE 1** *Laser-driven icon display, showing red, amber and amber/red icons*

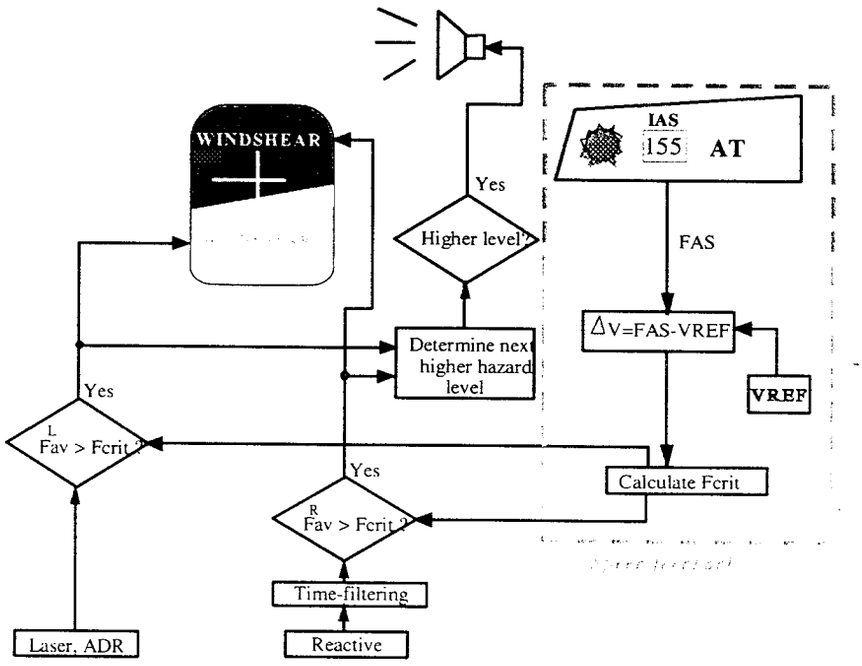


**FIGURE 2** *Integrated system for alerting and presenting labels, including speed feedback*

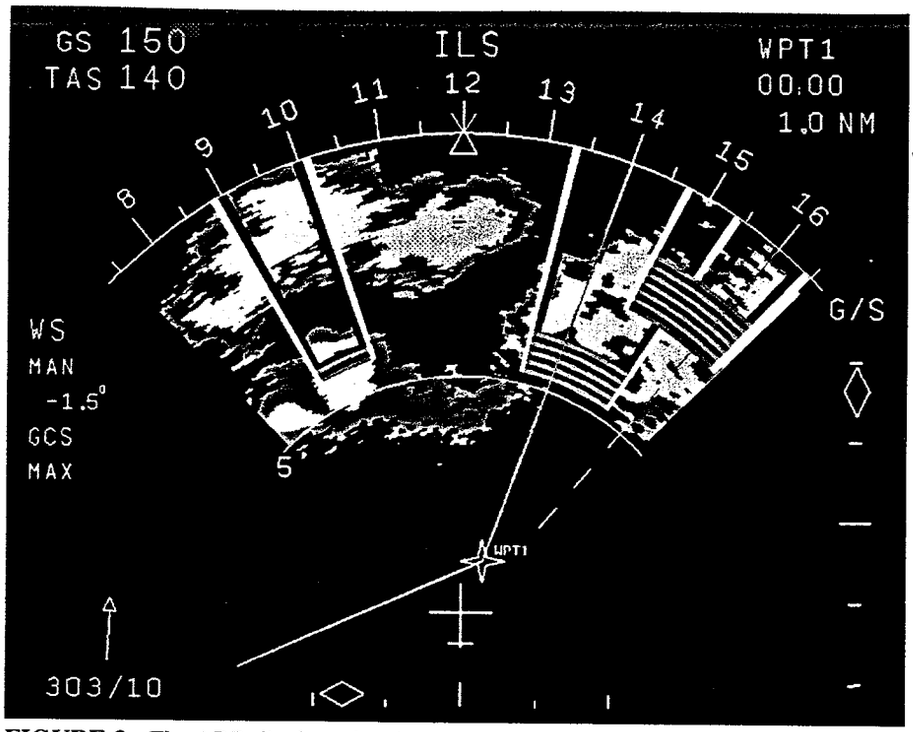


**FIGURE 3** *The ADR display, showing several windshear icons*

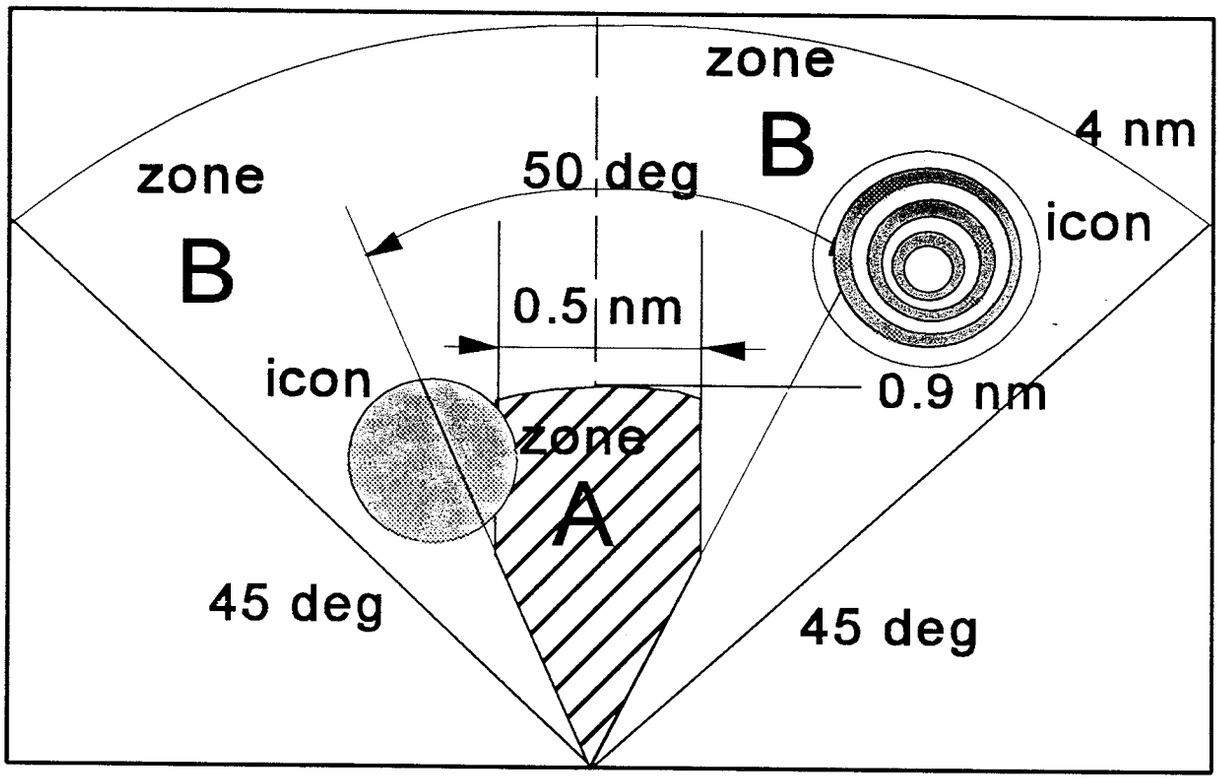


**FIGURE 4a** *Caution (B) and hazard zone (A) for the NAV display*

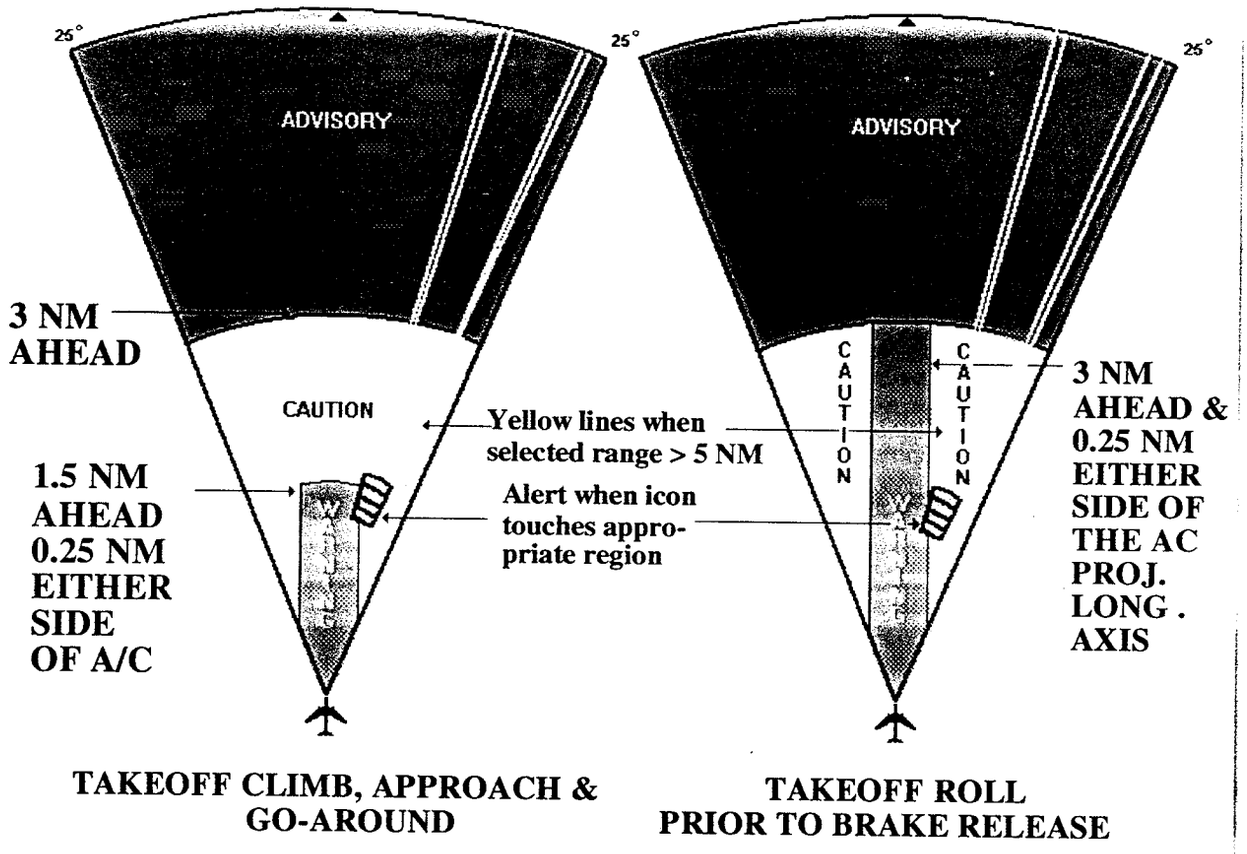


**FIGURE 4b** *Caution and hazard zone for the WX radar display, for take-off and approach (Ref.16)*

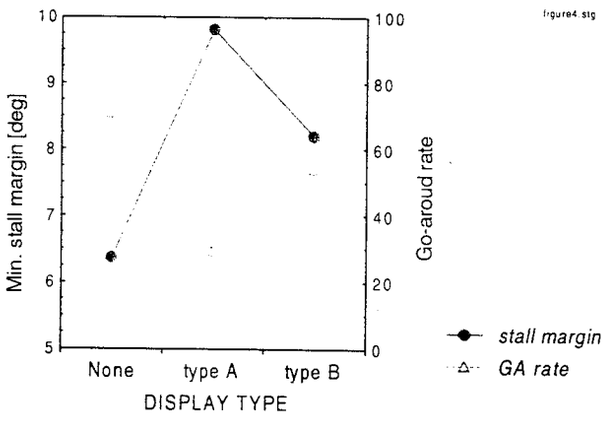


**FIGURE 5** *min. stall margin as function of display type*

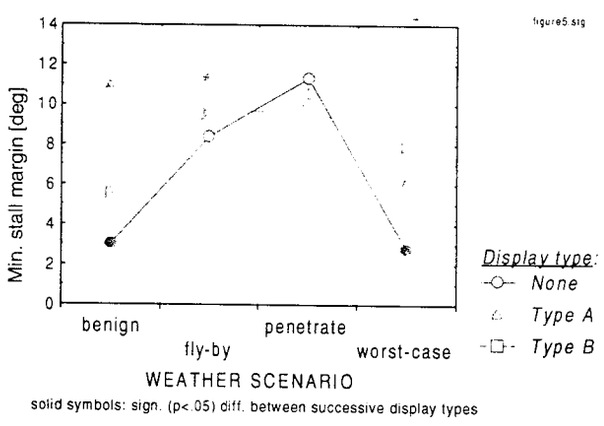


**FIGURE 6** *Min. stall margin as function of weather scenario, per icon display type*

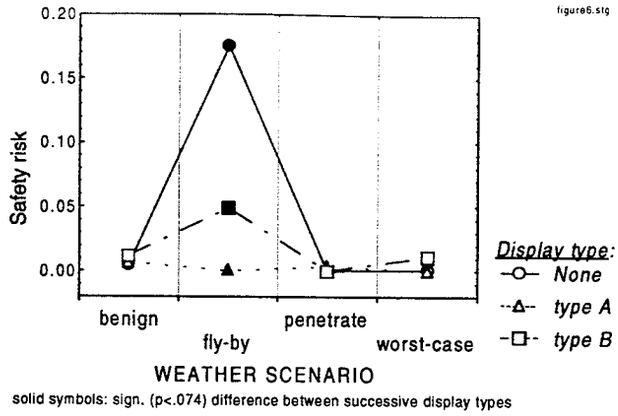


**FIGURE 7** *Safety risk as function of weather scenario, per icon display type*

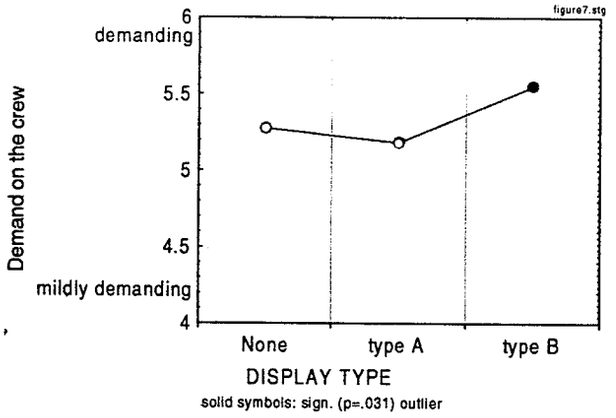


**FIGURE 8** *Demand on the crew as function of display type*

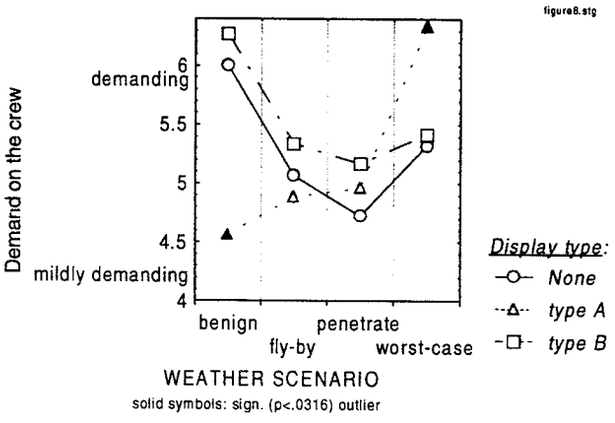


**FIGURE 9** *Demand on the crew as function of weather scenario, per icon display type*

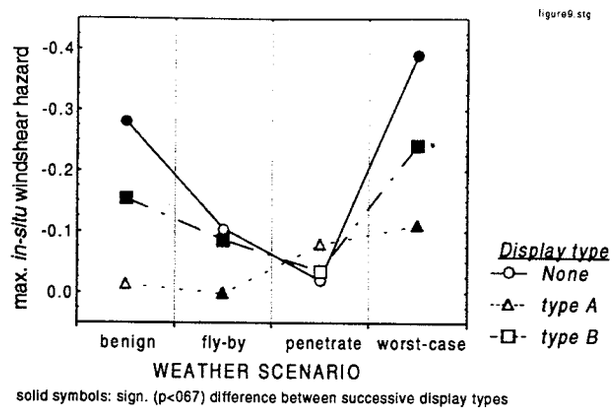


**FIGURE 10** *Effect of weather scenario, per display type, on the max. **in-situ** windshear hazard*

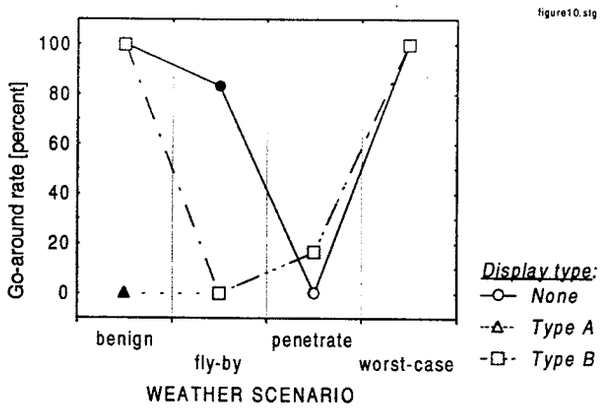


**FIGURE 11** *Go-around rate as function of weather scenario, per icon display type*

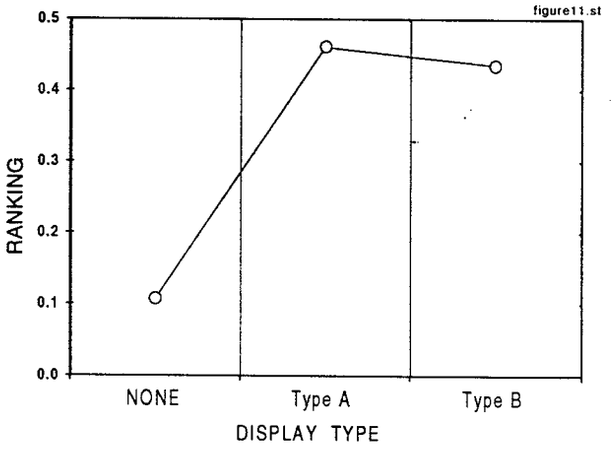


**FIGURE 12** *ranking of icon display according to pilot's preference*

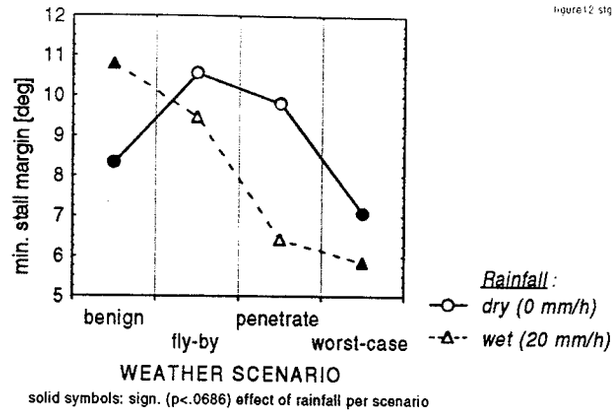


**FIGURE 13** *Effect of rainfall, per weather scenario, on the min. stall margin*

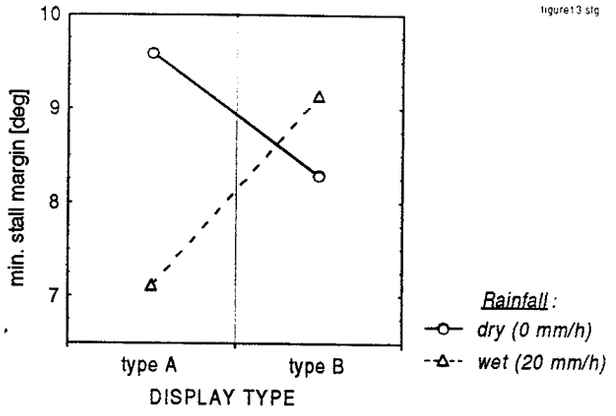


**FIGURE 14** *Effect of rainfall, per display type, on the min. stall margin*

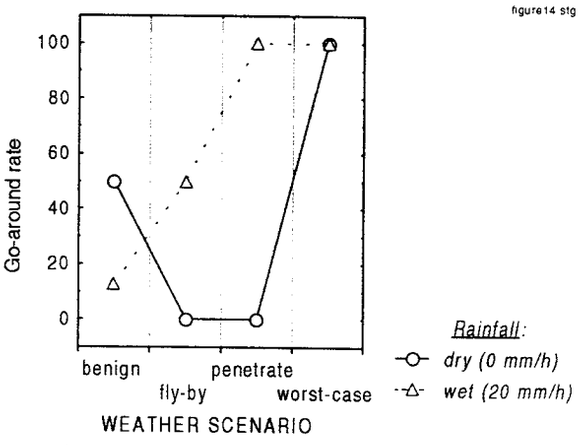


**FIGURE 15** *Effect of rainfall, per weather scenario, on the go-around rate*

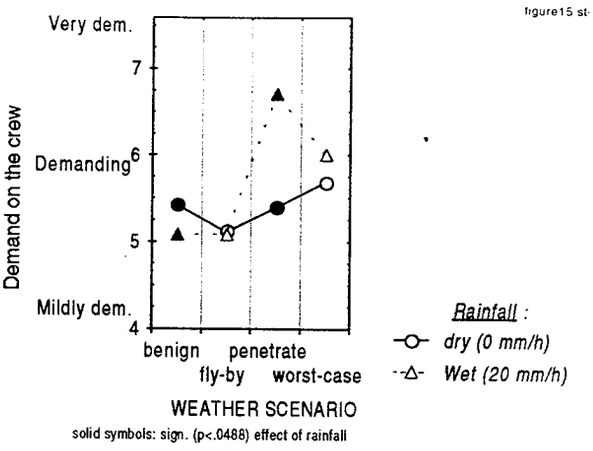


**FIGURE 16** *Effect of rainfall, per scenario, on the crew's workload*

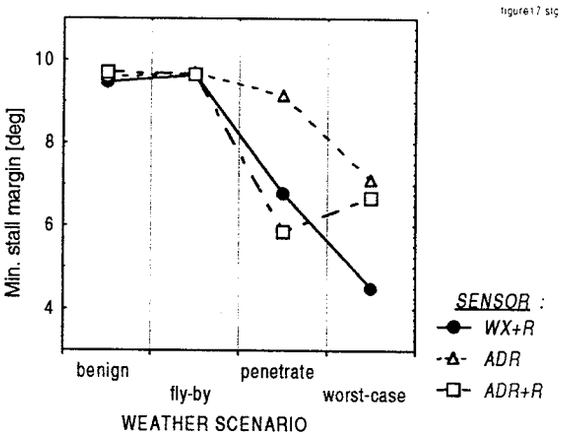


**FIGURE 17** *Effect of ADR sensor on min. stall margin, per weather scenario*

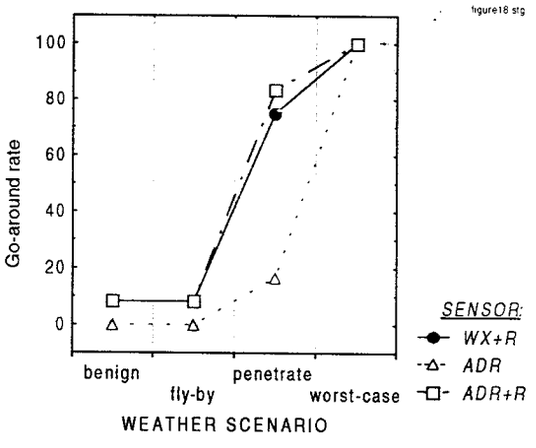


**FIGURE 18** *Effect of sensor, per weather scenario, on go-around rate*

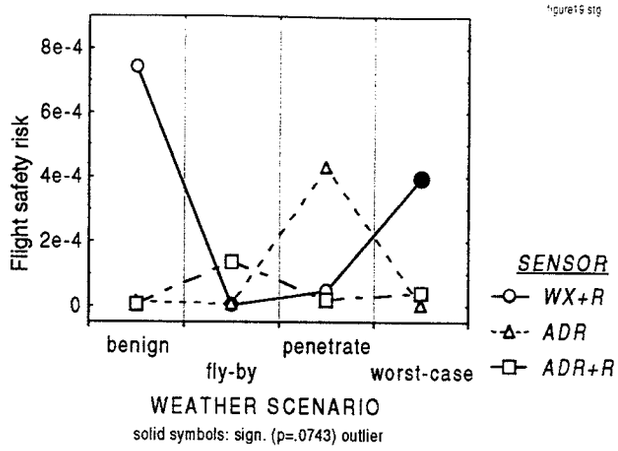


**FIGURE 19** *Flight safety risk as function of weather scenario, per sensor type*

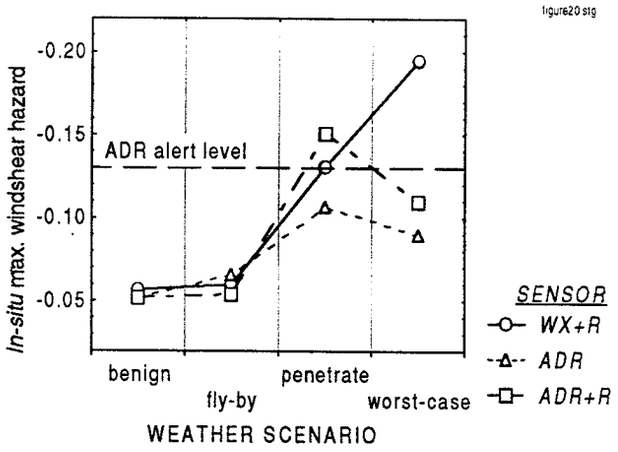


**FIGURE 20** *in-situ max. windshear hazard as function of weather scenario, per sensor type*

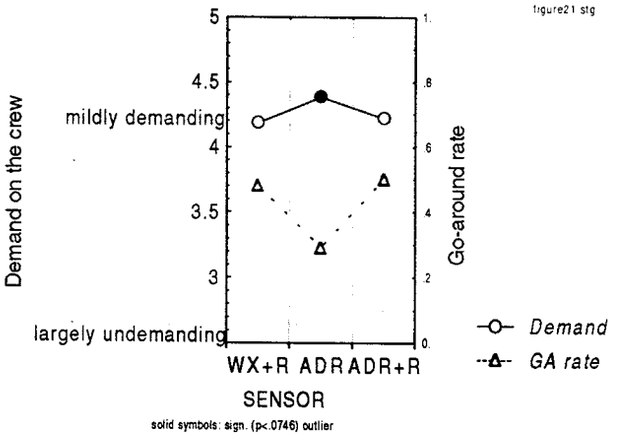


**FIGURE 21** *Effect of sensor on crew workload and go-around rate*

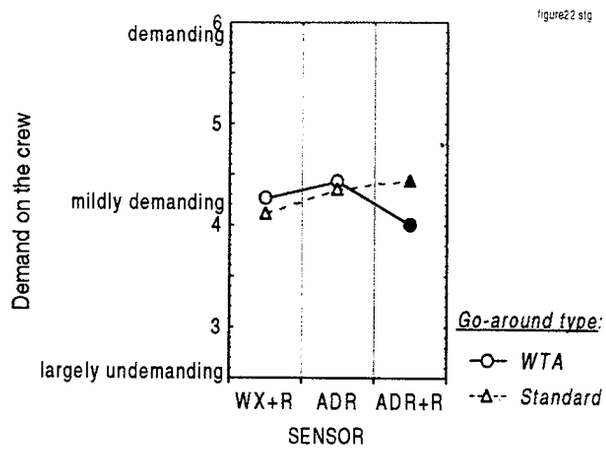


**FIGURE 22** *Effect of interaction between sensor and type of go-around on the crew's workload*

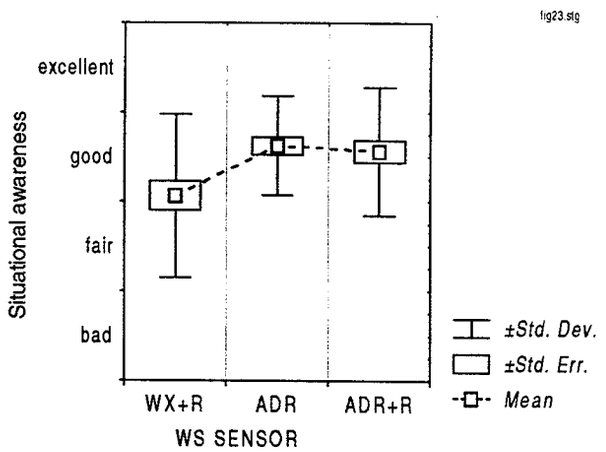


**FIGURE 23** *Effect of sensors on situational awareness, for the fly-by and worst-case scenarios*

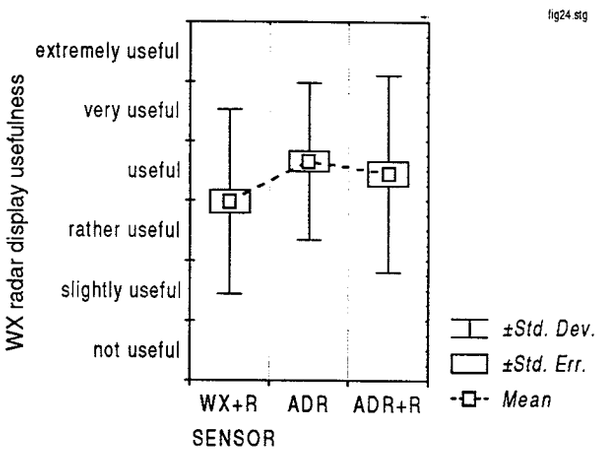


**FIGURE 24** *Effect of sensors, for fly-by and worst-case scenario, on WX radar's display usefulness*

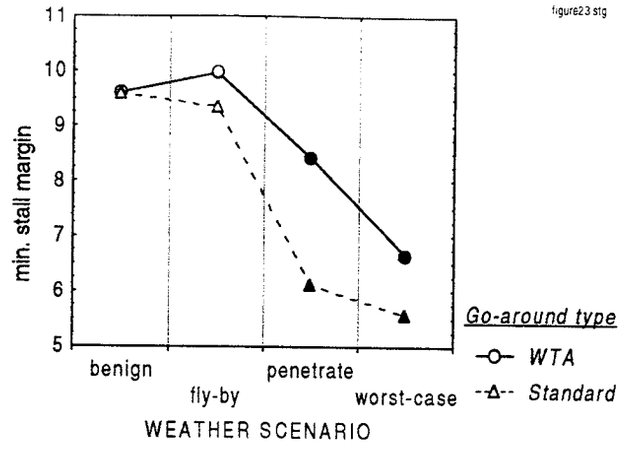


**FIGURE 25** *Effect of Go-around type, per weather scenario, on the min. stall margin*

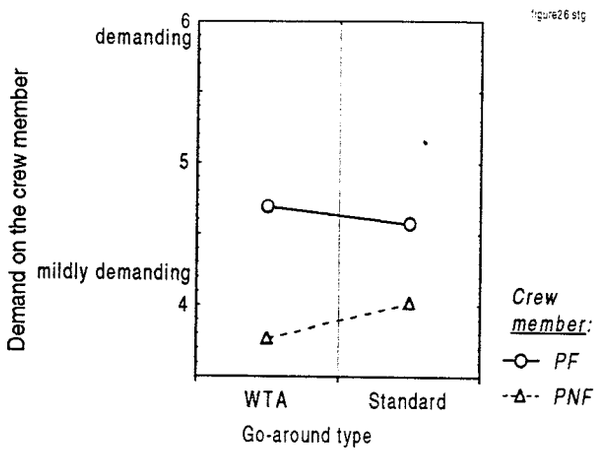


**FIGURE 28** *Effect of interaction between go-around type and crew member on the crew's workload*

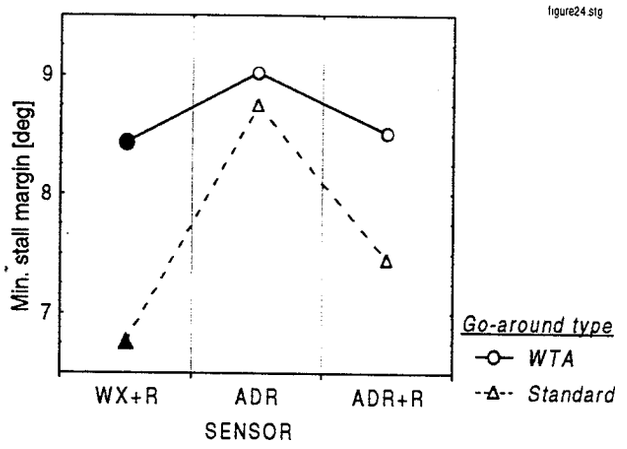


**FIGURE 26** *Effect of Go-around config., per sensor, on the min. stall margin*

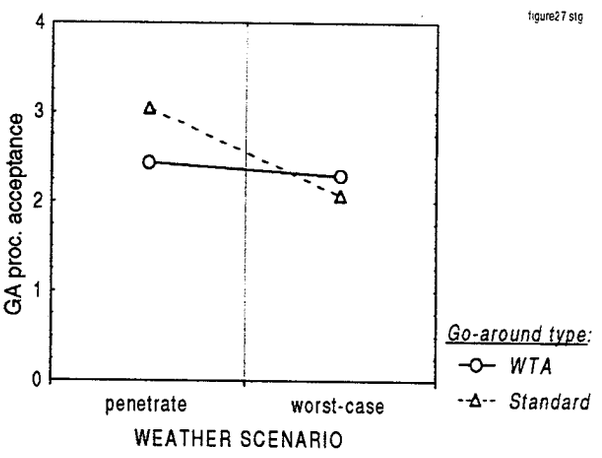


**FIGURE 29** *Acceptance of go-around procedure as function of weather scenario*

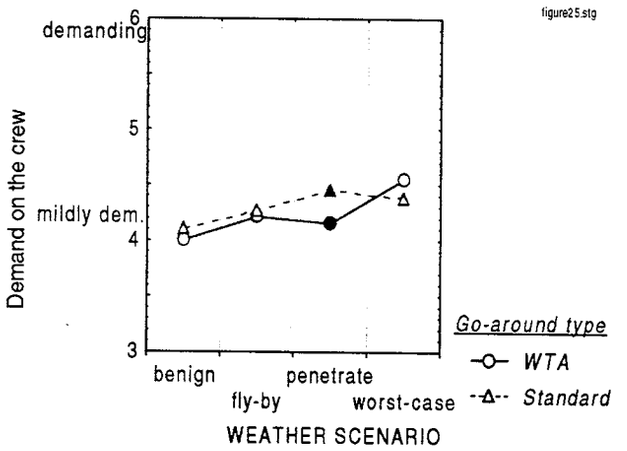


**FIGURE 27** *Effect of go-around type, per weather scenario, on the crew's workload*

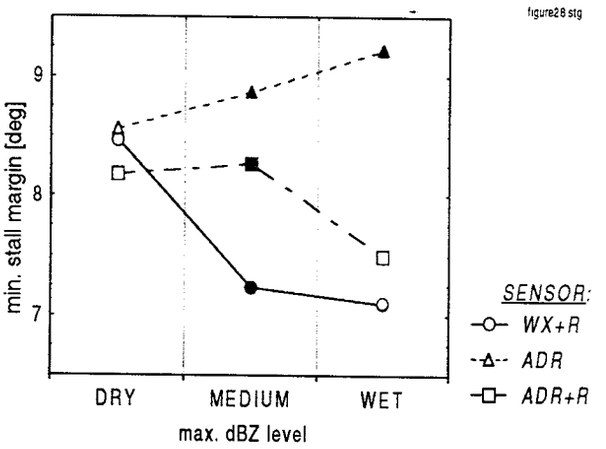


**FIGURE 30** *Effect of precipitation and sensors on min. stall margin*

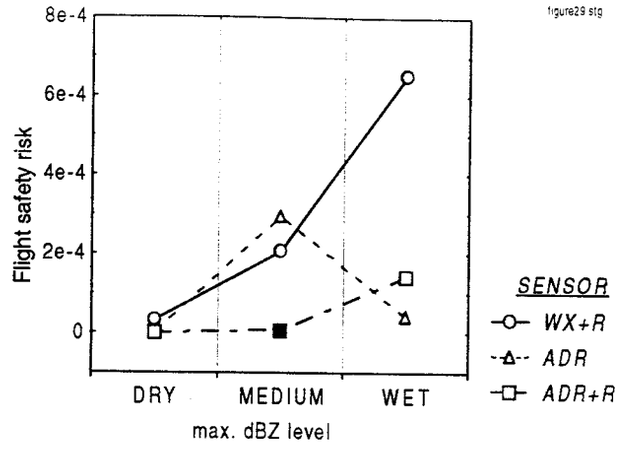


**FIGURE 31** *Effect of precipitation, per sensor, on the flight safety risk*

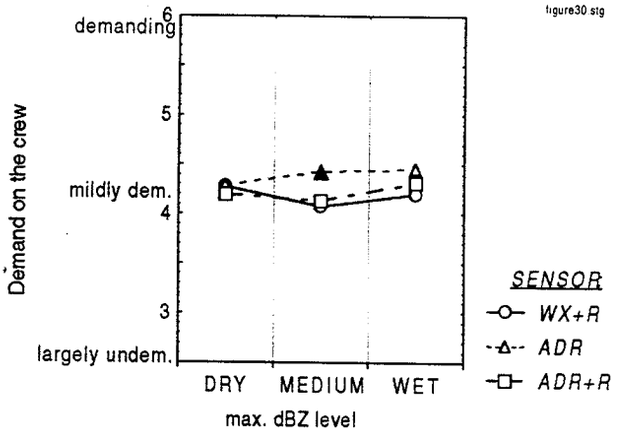


**FIGURE 32** *Effect of precipitation, per sensor type, on the crew's workload*

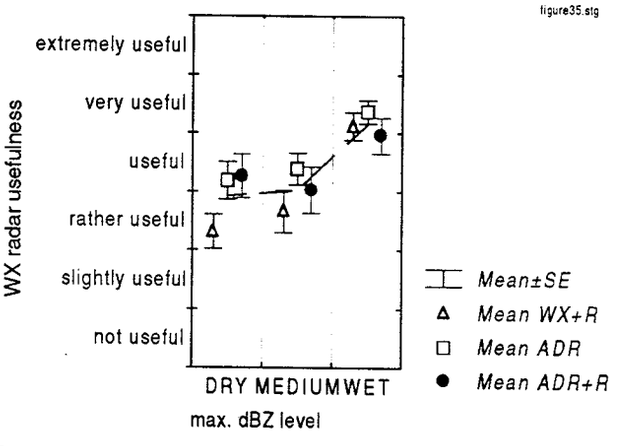


**FIGURE 33** *Usefulness of WX radar display as function of precipitation level, per sensor*

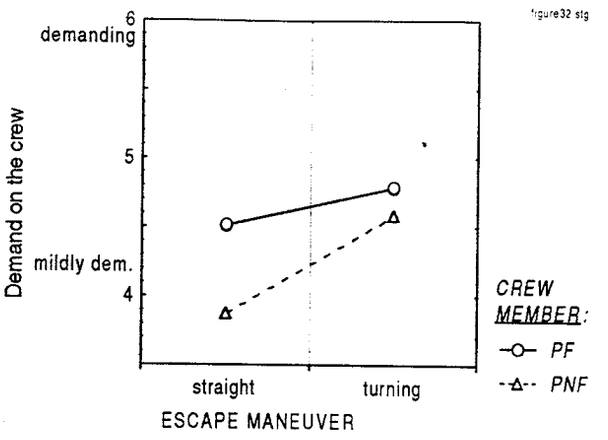


**FIGURE 34** *Effect of escape maneuver, per crew member, on the demand on the crew*

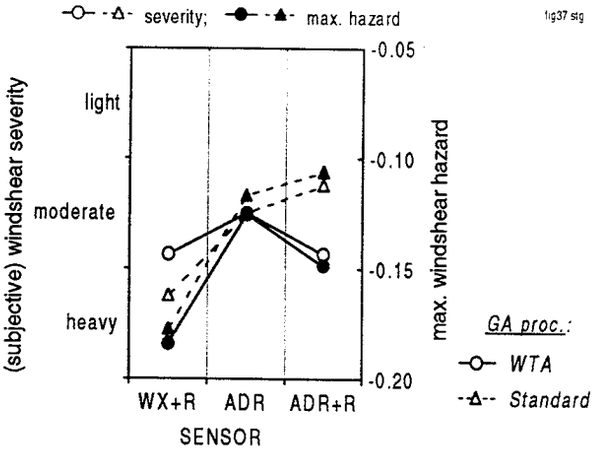


**FIGURE 35** *Effect of sensors and go-around procedures on subjective and objective windshear hazard*

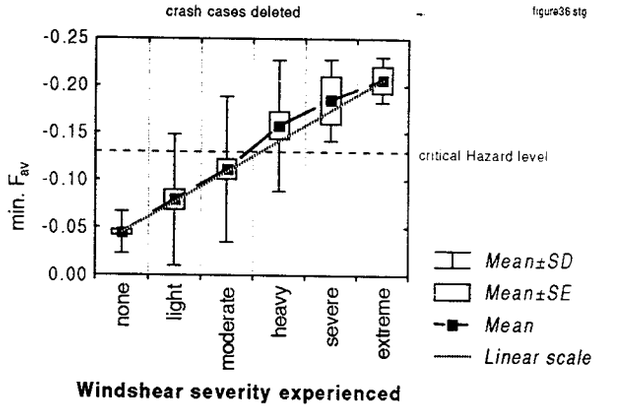


**FIGURE 36** *Relationship between subjective windshear severity rating and max. **in-situ** windshear hazard 'min. $F_{av}$'*