Assessing error tolerance in flight management systems

Hazel Courteney
CAA (UK) Safety Regulation Group

Abstract

Flight Management Systems (FMS) have been criticised for being too complex, and vulnerable to crew error. However, there have been little operational data available. This paper summarises results from a new study, recording FMS related events in service. Results suggest that the issues may be different from those first envisaged: for example, incompatibility between the aircraft system and the wider aviation environment seems more apparent than lack of understanding on the part of the crew.

In addition to reporting these data, suggestions will be made to address the issues raised, primarily from the regulators viewpoint. These will include 'human factors' acceptance criteria, evidence of competent development activity, in-service flaw reporting, and formal links to training programmes.

Views expressed in this paper are the opinion of the author and do not necessarily represent CAA policy.

Introduction and Summary

In contemporary aviation psychology, there has been much discussion about risks that might arise from pilot interaction with complex Flight Management Systems (FMS). The recent FAA Human Factors Task Team highlighted the complexity of FMS modes as a safety issue (Ref.1). Pilots have been quoted as asking 'What's it doing now?', 'Why did it do that?', or 'What will it do next?' (Ref. 2). Research papers have been published suggesting serious shortcomings in qualified pilots' understanding of their aircraft FMS systems (Ref. 3). However, there seems to have been little operational data to support these concerns, assess the prevalence of various kinds of events, or place them in context.

This paper will present new data obtained in thousands of operational sectors, showing that the predominant nature of events involving FMS are not necessarily those that the literature might have anticipated. Many of the pilots in this study were experienced FMS users. Their reports frequently reflected flaws in the system rather than incomplete understanding by the pilots, especially on aircraft Types that have been long established in service. Perhaps the cliché 'What's it doing now?' will, in the future, be replaced by 'Look what it's doing now', 'On no, it's doing *that* again!' or even 'We'll have to work around what it's going to do next'.

Results of Data Collection Study

Three major UK operators collected data on operational events involving the FMS. This spanned eight fleets, recording a total 2066 sectors, of which 1134 had 'no event'; the remaining 932 logged 1779 events. Fig. 1 shows data on relative incidence of some selected event classes for the main aircraft Types. All event classes for all aircraft are illustrated in Fig.2.

Fig.1 Relative Frequency of Selected Events by Aircraft Type

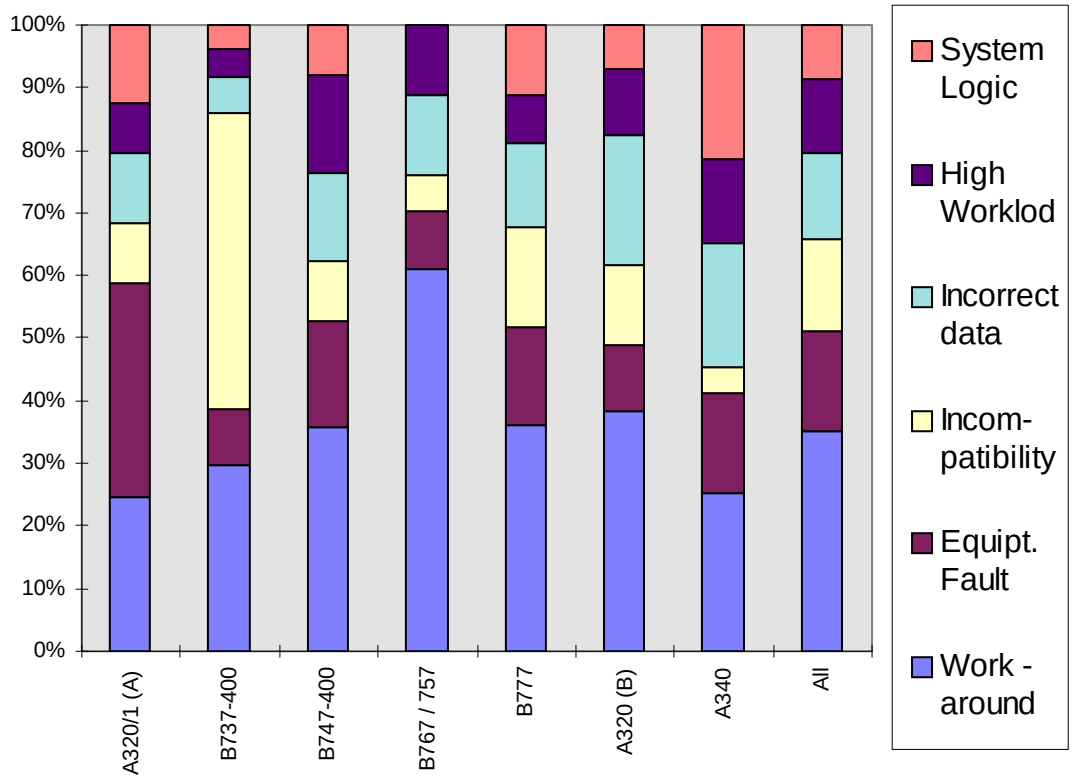
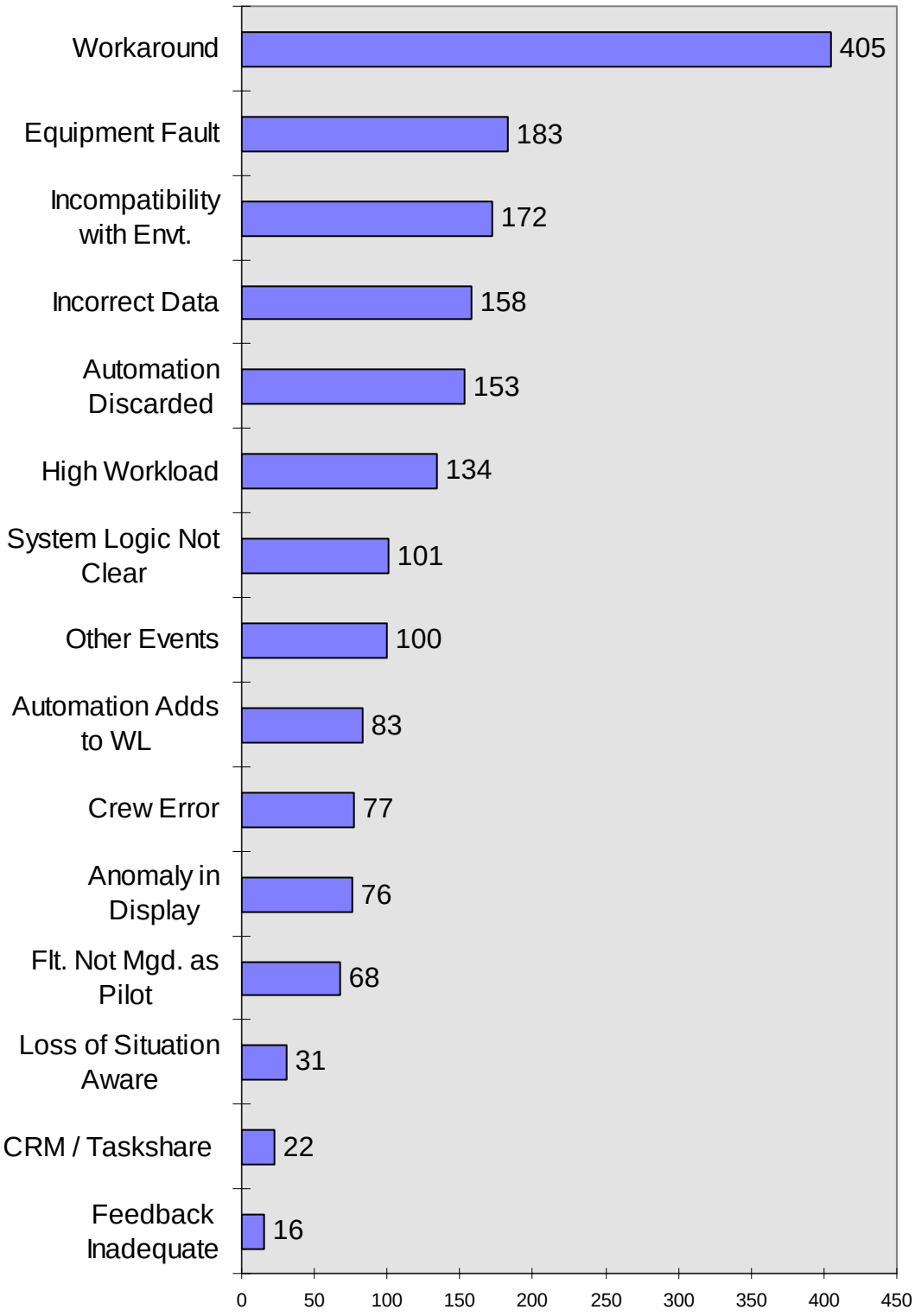


Fig. 2 Events Reported In Service



Implications of Data

The nature of the reported events is interesting and not entirely predicted by the literature. The anticipated flood of reports of 'inadequate feedback', 'crew error', 'reduced situation awareness', or 'system logic not clear' did not appear. It could be argued that this is influenced by the self reported nature of the data, but it is difficult to envisage other methods that would be more accurate. However, comments on the study indicated that the pilots usually felt quite confident in their understanding of the system that they were using (although some indicated that this might not be equally true with less experienced pilots, or crew serving with other operators). What these data suggest is that:

- Validation of FMS design requirements against user experience and the ATC environment is incomplete.
- Validation of the FMS design functionality and data against known requirements is incomplete.
- Existing training is not necessarily sufficient to ensure the crew can predict aircraft behaviour at all times.
- The system design must assume that crews will make routine errors and accommodate this characteristic.
- Procedures should be compatible with design intent, to avoid problems such as late (and unachievable) reprogramming.
- Flaws and difficulties with the system can persist in service, because they are not always - or even usually - dealt with effectively.
- Safety implications may arise from indirect aspects such as increased head down time, erosion of manual flying skills and late runway changes requiring rushed reprogramming.

How Can These Issues be Addressed?

The FMS, and its integration into the greater civil aviation environment, is a large and complex system, and probably almost impossible to fully assess as a completed design. For the future, comprehensive 'mature' requirements for system certification or approval should cover a timespan that extends back through development, and forward through in-service use.

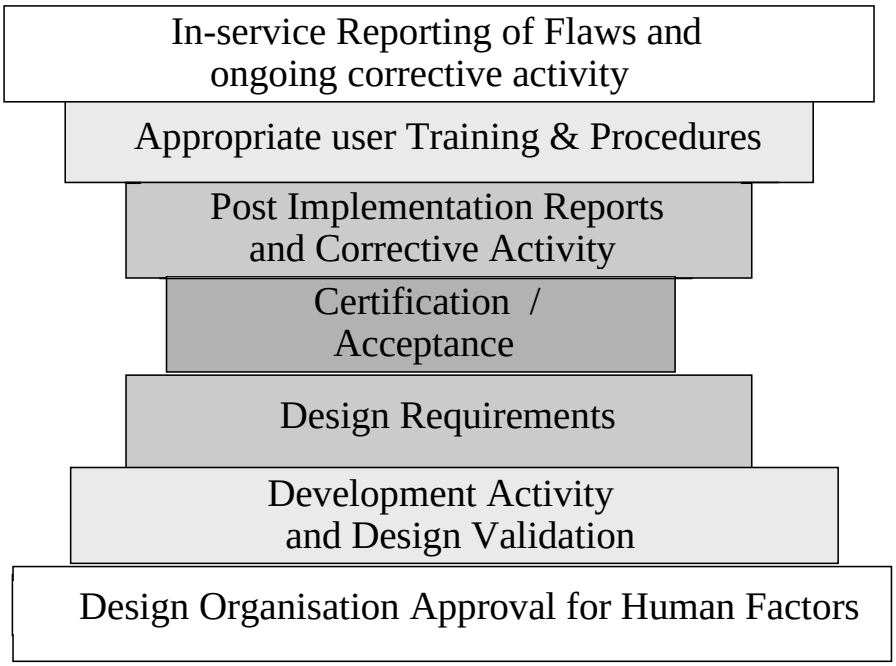


Fig. 3: Hourglass Model: Seven Stages of Requirement

This model of 'mature' requirements is shaped like an hourglass (Fig 3). The narrow part in the centre brings the product into sharp focus and accepts or rejects it against specific, predetermined criteria. The stages before and after this event are more flexible, general and 'diffuse' with increasing distance from the product acceptance point. Starting from the bottom up, the stages describe building confidence from the most basic foundation, through development and into service.

i) *Design Organisation Approval:* The first (bottom) level in Fig. 3, refers to the overall assessment of an organisation (by a customer or regulatory authority) for general competence as a supplier or designer of aerospace equipment. In the future, this could include 'maturity' of organisational competence in human factors activities (example, Ref. 4).

ii) *Development Activity:* The second level covers the development process as specifically applied to the product in question. The precise nature of the activities would depend upon the particular design requirements, and the means selected by the manufacturer to demonstrate their compliance. It is very likely that an acceptable means of compliance would include, as a minimum:

- iterative structured evaluation of representations or prototypes, with test crews and representative end users (Ref.5). This must occur at stages of development where recommendations for

change can still be accepted. Modification activities should be planned into the program and adequately resourced.

- a comprehensive approach to initial and continuing integration of the system with other features of the environment in which it will operate (such as Air Traffic and aerodrome procedures, training and procedural expectations, maintenance capability).
- the application of systematic technique(s) for the identification of potential risks arising from inappropriate human actions (including flight crew and maintenance personnel).
- a review of known risks to appropriate class of aircraft operations, and evaluation of implications that the new system might have (direct and indirect) for such risks. For example, according to the current CAA SRG Business Plan the top safety risks for UK aviation is Crew and Human Factors, including:
  - *Omission of action / inappropriate action*
  - *Flight Handling*
  - *Poor professional judgement / airmanship*
  - *Failure in Crew Resource Management* (*CRM)*
  - *Lack of positional awareness in the air*
  - *Maintenance human factors*

  Could an FMS implementation have adverse effects in terms of an omitted or inappropriate action; Flight Handling (due to situations developing during head down time); position awareness (e.g. due to incorrect data); or a breakdown in CRM (Ref.6)?

iii) *Design Requirements:* The third level refers to the specific product requirement that would appear in a regulation, a commercial contract, or a requirement specification. This would include criteria to ensure realistic integration of the human user into the system. There are various ways to frame such requirements; the means favoured by the author is described as ‘abstract prescription’. This style of criteria is relatively objective, and clear in discriminating between a design feature that is, or is not, acceptable, but it does not restrict the choice of design solutions unnecessarily. The following examples for ‘error tolerance’ are offered, not necessarily as recommended criteria, or as a complete list, but simply to illustrate the style:

- 'a single slip or lapse (e.g. incorrect data entry) should not be capable of progress to a potentially hazardous outcome without direct and compelling feedback to the crew'

- 'no single crew action with a hazardous consequence (e.g. in an aircraft, closing down the only remaining propulsion, or all hydraulic power) should be possible without a direct challenge from the system that must be positively overridden by the crew. Such challenges should not be presented when the action is not hazardous, to prevent it from becoming perceived as routine.'

iv) *Certification, approval or acceptance* must occur at a defined point in time, when the design is complete, and before it enters service. The decision to accept the design can be supported by evidence of organisational competence and thorough development activities. It is suggested that the certification stage should, in the future, also take account of the adequacy of planned *subsequent* stages of the product life cycle (v, vi, and vii, described below).

v) *Post Implementation Reporting (PIR):* Where a system is complex, it seems likely that not every detail can be guaranteed correct before implementation. This is surely true of one (such as FMS) that has interfaces with many external aspects, including human crew, airport procedures world-wide, and a dynamic international air traffic system. It therefore seems reasonable that approval of such a system should include a requirement for an effective post implementation user reporting system.

vi) *Training and operating procedures* for system users have generally been considered as issues that are separate from system design. This being the case, the design can be approved, or otherwise, independent of the training and procedures that will be available for system users. However, this presents a difficulty, because in human factors terms, a design may sometimes be acceptable if - and only if - an acceptable degree of training is provided, and a minimum level of competence is achieved. In the future, it may become necessary to directly link issues raised during the development and approval of systems with the minimum training and achievement that will accompany it. For aircraft, this could potentially be addressed through the Type Rating requirements.

vii) *In-service reporting:* The data gathered in the FMS study indicate that there are numerous imperfections operating in service that are known to line crews. Although there are systems in existence that purport to pass back in-service problems to manufacturers, respondents in the present study confirmed that *none* of the events in this study were reported through other avenues. This leads to the conclusion that approval of a new system could beneficially include a requirement for effective in-service reporting mechanisms that will identify the myriad of minor - and not so minor - imperfections, for ongoing corrective activity aimed at updates and

subsequent products. This would be a lower level, less intense version of the PIR described in v) above.

Conclusions

The data on FMS operations in service suggests that there are human factors issues beyond mode complexity. This includes crew being distracted by incompatibility between the FMS design and the operating environment, incorrect data and anomalies in the system, as well as training and procedures that are not sufficient for comprehensive system utilisation. It is suggested that, in order to address these issues, there should be requirements for:

- better validation of the system requirements and the design
- a much improved mechanism for in-service feedback
- direct links between the system features and the training and procedures that accompany it into service.

References

1. FAA Human Factors Task Team Report on: 'The Interfaces Between Flightcrews and Modern Flight Deck Design', 1996
2. James, M., Birch, C., McClumpha, A., Belyavin, A.: 'The perception of workload on automated flight deck', 1993
3. Sarter, N. B., and Woods, D. D., 'How in the World Did We Ever Get into That Mode? Error Awareness and Supervisory Control', in Human Factors, 1995, 37(1), 5-19
4. Earthy, J.V., & Tomlinson, C., 'Human-Centred Process Assessment: A Model and Maturity Scale', written under the INUSE project, D5.1.7(m) 1997
5. Courteney, H.Y., 'User Error: Practical Ways to Protect Your System' in the proceedings of the Conference of the International Council of Systems Engineers (INCOSE), Vancouver 1998
6. Plat, M., & Amalberti, R.: 'Experimental Crew Training to Deal With Automation Surprises'.: IMASSA, Département Sciences Cognitives Brétigny-sur-Orge, 91223-F, France