Performance Monitoring and Assessment of Neuro-Adaptive Controllers for Aerospace Applications Using a Bayesian Approach

Pramod Gupta[*], Kurt Guenther[†], John Hodgkinson[‡] Stephen Jacklin[§]
Michael Richard[**], Johann Schumann[††], and Fola Soares[‡‡]

Abstract

Modern exploration missions require modern control systems-control systems that can handle catastrophic changes in the system's behavior, compensate for slow deterioration in sustained operations, and support fast system ID. Adaptive controllers, based upon Neural Networks (NN) have these capabilities, but they can only be used safely if proper verification & validation (V&V) can be done. In this paper we present our V & V approach and simulation result within NASA's Intelligent Flight Control Systems (IFCS).

I. Introduction

Modern aircraft, UAVs, and robotic spacecraft pose substantial requirements on controllers in the light of ever increasing demands for reusability, affordability, and reliability. The individual systems (which are often nonlinear) must be controlled safely and reliably in environments where it is virtually impossible to analyze—ahead of time—all the important and possible scenarios and environmental factors. For example, system components (e.g., gyros, bearings of reaction wheels, valves) may deteriorate or break during autonomous UAV operation or long-lasting space missions, leading to a sudden, drastic change in the vehicle performance. Manual repair or replacement is not an option in such cases. Instead, the system must be able to cope with equipment failure and deterioration. Controllability of the system must be retained as good as possible or re-established as fast as possible with a minimum of deactivation or shutdown time. In such situations the control engineer has to employ adaptive control systems that automatically sense and correct themselves whenever drastic disturbances and/or severe changes in the plant or environment occur.

Over the recent years, artificial neural networks (NN) have found their way into various safety-related and safety-critical areas, like transportation, avionics, environmental monitoring/control, and medical applications. Although many of these applications have turned out to be highly successful, they also pose high risks and significant development costs, producing concerns and reluctance to adopting these new and sometimes complex and difficult-to-understand technologies. Of chief concern is the general question of how can it be guaranteed that the NN-based adaptive control system performs as expected. While theory and concepts of adaptive systems and intelligent control have been studied in depth over the past decade or so, only very little attention has been paid to the issue of validating the correctness and guaranteeing safety of their operation and monitoring their performance during operation.

Validating the correct performance of a controller requires a set of concise design requirements and performance criteria. In the case of control systems for piloted aircraft, generally applicable quantitative design criteria are very difficult to obtain. The reason for this is that the ultimate evaluation of a human-operator control system is necessarily subjective and, with aircraft, the pilot evaluates the aircraft in different ways depending on the type of aircraft and phase of flight. In most aerospace applications (e.g., for flight control systems), performance assessment is carried out in terms of *handling qualities*. Handling qualities may be defined as those dynamic and static properties of a vehicle that permit the pilot to fully exploit its performance and other potential in a variety of missions and roles. Traditionally, handling quality is measured using the Cooper-Harper rating done subjectively by human pilots for aircraft control. In our study, which will be described in this paper, we use a quantitative approach

[*] Computer Scientist, QSS/NASA Ames, M/S 269-3,.
[†] AS&M/NASA DFRC .
[‡] Spiral Technology/NASA DFRC.
[§] Computer Scientist, NASA Ames, M/S 269-2.
[**] Aerospace Engineer, NASA DFRC.
[††] Senior Scientist, RIACS/NASA Ames, M/S 269-3. .
[‡‡] Contek Research, El Sugando, CA

using low order equivalent system (LOES) model of the aircraft. The LOES approach is to match the high order responses with the low order responses of the familiar modes. LOES allows specification of model dynamics consisting of low order systems with preferred values of damping, frequency or time constant. A central element for assessing the performance of a neuro-adaptive controller is the ability to dynamically monitor the performance of the adaptive neural network.

We have developed a set of related tools, which can be used in all phases of the software lifecycle (including system design, simulation, system prototype development, deployment, and in-operation monitoring) to assess the performance of neural controllers *while in-flight.* These tools (confidence tool, envelope tool) are based on Bayesian methods and are capable produce statistical confidence intervals for the controller signals. Using this knowledge of the error and of the network or model, our tools will allow real-time assessment of vehicle performance and provide an estimate of important handling quality characteristics.

In this paper we will present the methods of measuring the performance of a neuro-adaptive controller and how the tool performance metric relates to the control system robustness and the vehicle handling characteristics. Simulation results will be presented and tool design concepts will be explained.

II. IFCS Control Architecture

We will illustrate our approach with the adaptive flight control system (FCS), which has been developed within the IFCS project at NASA. The target aircraft for this controller is a specifically configured F15 jet aircraft. It is equipped with additional actuator surfaces (canards) that are located in front of the wings. By moving them, the airflow over the wing can be modified in a wide range. Thus, this aircraft can be used to simulate failures like wing-damage during test flights. The FCS (Figure 1) is a straight-forward dynamic inverse controller: the pilot steering commands are mixed with the current sensor readings (airspeed, angle of attack, altitude) to form the desired behavior of the aircraft (measured as roll-rate, pitch-rate, and yaw-rate). The dynamic inverse model then calculates the required actuator movements (e.g., of aileron or rudder) to bring the aircraft into the desired state. If the aerodynamics of the aircraft changes (e.g., due to a broken surface), there is a deviation between desired and actual state. The neural network is trained during operation to produce a correction signal U_{AD} to minimize this deviation. .

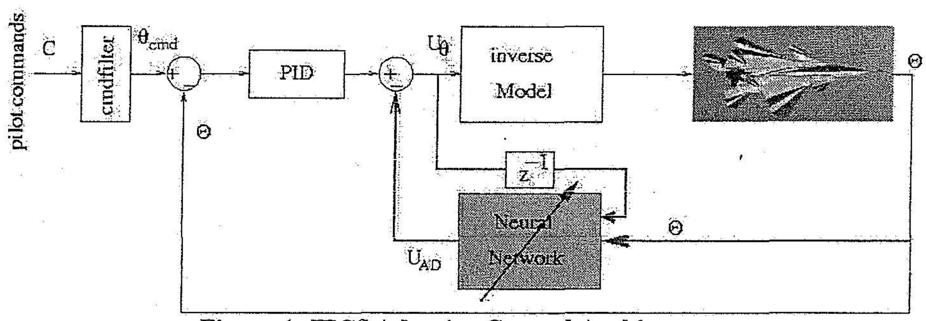

Figure 1: IFCS Adaptive Control Architecture

The controller (IFCS Gen-II, [4]) uses a Sigma-Pi neural network. In this type of neural network, the inputs are subjected to basis functions (e.g., square, scaling, logistic function). The output of the network is a weighted sum of a Cartesian product of these function values. For details of the control architecture and the network learning rule[1,7].

III. Simulation Environment

The experiments have been carried out on the Dryden F-15 simulator. The Dryden F-15 simulator is a 6-DOF Real-time simulation that simulates flight of the F-15B and the F-15 ACTIVE aircraft. The Real-time complex consists of a Sun computer, a cockpit and a projector system for out the window (OTW) graphics. The complex has the ability to interface to external hardware for HIL testing. The cockpit has its own embedded computers for the electric stick and cockpit interface unit. These components communicate with the main Real-time computer via high speed data bus. The Sun computer is a Unix-based system with multiple CPUs. The machine is equipped with 32

GB of memory and over 200 GB of disk space. The main simulation models are programmed in FORTRAN-77, C/C++, Ada-95 and openGL. These simulation programs simulate the actions of the F-15's aerodynamics, FCS, engines and other subsystems with a basic frequency. The entire complex is controlled via a Sun console where the user can enter commands to control the simulation and is operated by two people; one to fly and the other to operate the console.

IV. □ Network Performance Measuring and Assurance

One of the key factors which limit the use of neuro-adaptive systems in safety critical applications has been the difficulty of demonstrating that the neural network will generate reliable outputs once it is in routine use. We have no way of quantifying the magnitude of errors one encounters in the output of the network. For this reason the method of confidence interval can be used: a confidence interval gives an indication of how much uncertainty there is in our estimate of output; the narrower the interval, the more precise is our estimate.We have developed the Confidence tool, which is based on a Bayesian statistical model of the neural network. The tool considers the probability distribution of the NN output, based upon the posterior distribution of the weights. The confidence tool dynamically calculates the current performance characteristics of the system and thus provides a dynamic measure of how reliable the current approximation of the system is. The details of the tool can be found in[9,10]

Figure 2 shows simulation results.. The experiments were conducted at a flight condition: 20,000 feet altitude and 0.75 mach. Starting in a nominal mode, a failure (canard failure) is occurring after some time. Around this time, the pilot initiates a sequence of doublet inputs, which makes the failure apparent and causes the NN to adapt. The blue line in Figure 2 shows the output of the neural network (control augmentation signal Uad); the red lines comprise the confidence intervals around the network input. A narrow interval at the beginning shows that the network has (as expected) a high confidence for the nominal mode. After the failure occurs, the confidence is dropping substantially (large confidence interval)-an indication that the network still has to adapt toward handling the failure scenario. Subsequent pilot commands still causes network adaptation. However, the confidence interval is much smaller now, a clear indication that the network is successfully learning to cope with this situation.

V.□ Handling Quality Analysis

5.1 Handling Quality Metrics: Cooper Harper Rating and MIL 1797

Handling qualities are defined as "those qualities or characteristics of an aircraft that govern the ease and precision with which a pilot is able to perform the tasks required in support of an aircraft role" (1). Pilots are asked to quantify handling qualities by assigning a numerical rating to a specific piloting task. The most common rating scale is the Cooper-Harper scale (see Figure 3), where numbers are awarded to different piloting situations according to their ease and precision[5].

This rating scale is a 10-point scale in which 1 is excellent and 10 represents actual loss of control. Three coarser levels of "Satisfactory" (rating 1-3), "Acceptable" (4-6), and "Controllable" (6-9) are often used instead. This rating is highly subjective, and thus difficult to describe in quantitative terms. The Military Specification MIL-F-8785 versions A, B and C (c. 1954-1975), defines quantitative criteria for different classes of aircraft (e.g., fighter, transport) and for different flight phases (Category A,B,C; e.g., combat, cruise, landing/takeoff). For an aircraft of a certain class, the MIL 1797 specifications require Level 1 handling qualities for flight conditions within the Operational Flight Envelope. In the Service Flight Envelope, Level 2 handling qualities are allowed. In a more extreme Permissible Flight Envelope, even Level 3 handling qualities are allowed.

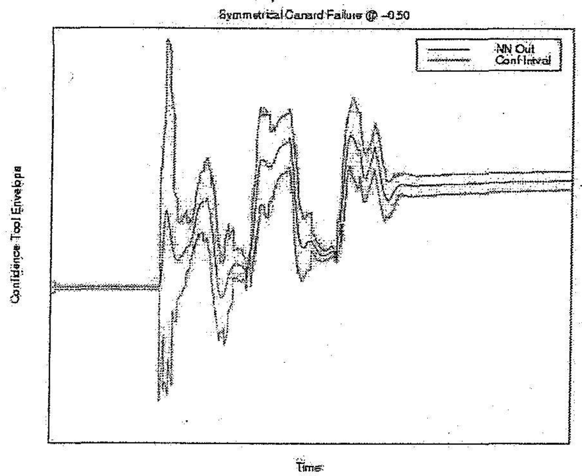

Figure 2. Simulation results for Canard failure. The blue line is the network output; the red lines mark the confidence interval.

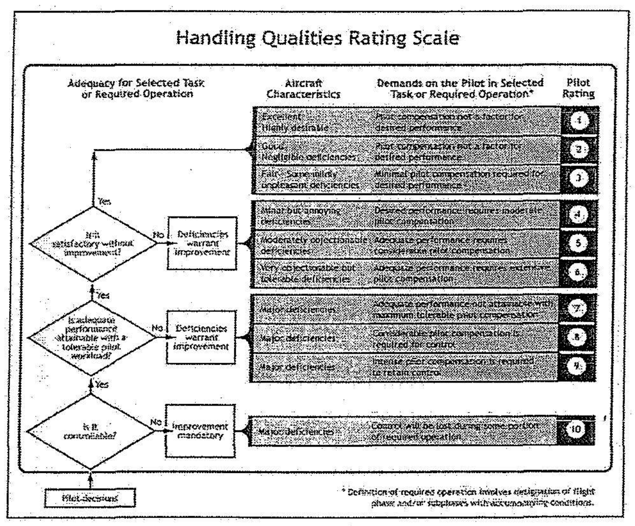

Figure 3. Cooper-Harper rating scale

Evaluating the baseline handling qualities of an adaptive system poses unique challenges to the V&V practitioners. First, the response criteria for failures during and after adaptation must be defined and then it must be determined whether an adaptive control system is producing adequate handling qualities during and after a failure. The criteria can be divided into three classes, i.e. failure transient criteria, primary handling criteria and secondary handling criteria. The first type governs the transient excursions as the failure happens and while the adaptation is still taking place. The second type governs the handling qualities of the aircraft after the failure and after the adaptation when the neural net has determined the set of weights best suited to the failed system. The second type is subdivided into two subtypes. The first subtype governs the conventional handling qualities of pitch due to pitch command, roll to roll command, yaw to yaw command, etc. and the second the off-axis handling qualities- for example, a failure in one side of a rolling tail may produce objectionable bank excursions in response to a pitch command.

5.2 Low order equivalent (LOE) Analysis

Standard methods of analysis were used to examine the predicted handling qualities including Control Anticipation Parameter (CAP) plots and bandwidth plots. Simulated flight data was analyzed using the lower order equivalent systems method to determine aircraft response parameters. A simple short period pitch response model was chosen, based on a lower-order model plus a delay term representing the longitudinal pitch axis. A state space format was used to present pitch dynamics that demonstrated Level 1 handling qualities. The simulations were conducted at the flight condition of Mach 0.75 at 20K, to approximate the standard IFCS Flight Envelope.

The lower order longitudinal model is defined by the following equation:

$$\frac{q}{\delta ep} = \frac{K_{lon}\,(s + L_{\alpha})\,e^{\tau s}}{s^2 + 2\zeta_{sp}\omega_{sp}s + \omega_{sp}^2} \qquad \left(\frac{\text{rad/sec}}{\text{inch}}\right)$$

where q is pitch rate, dep is pitch stick position, K_{lon} is the gain, w_{sp} is short period frequency, L_{α} is the dimensional lift curve slope, $e^{-\tau s}$ is a time delay term to approximate high frequency lag accrual and ξ_{sp} is short period damping.

A Fast Fourier Transform (FFT) technique was used on the time history data to obtain the pitch-rate-to-stick-input transfer function. The FFT used stick displacement as input, and pitch rate in deg/sec as output. The aircraft's parameters were estimated directly from the time histories of an in-flight pitch frequency sweep for the F-15 during a Gen I flight, yielding a pitch response that was subsequently reduced to frequency response data using FFT analysis. These data were in turn matched by hand with a short period transfer function giving the results in Figure 4. In this graph, results for flight data is shown in green. The system response for the LOE system is displayed in blue. The match of the aircraft and LOE behavior is adequate for the short period frequency range and the response being matched contains the airplane plus any closed- and open-loop control functions.

Figure 5 shows the Control Anticipation Parameter (CAP) for the above simulation experiment. The CAP summarizes the LOE response with respect to short period damping. According to MIL-1797, the CAP has to satisfy certain requirements for a specific level of handling quality (level 1 or 2). These requirements are depicted in Figure 5 as rectangular areas (bounded by red lines). In our case, the CAP (light blue dot at approx. (0.1,0.8)) is clearly within the Level 1 area, indicating a good short period damping performance.

5.3 Full Model Handling Quality Analysis

In the case of a nominal aircraft, LOE analysis, as discussed above is sufficient to determine the handling quality levels. However, damage in aircraft and non-linear effects, introduced by the adaptation, requires the use of an analysis method, which uses the full model of the aircraft and controller. In this paper, we describe two sets of experiments, which have been carried out using a Simulink model of the IFCS system. The detailed analysis of time delays and transients in pitch response comprises an important aspect of handling quality. Larger time delays can lead to uncontrollable situations due to pilot-induced oscillations. We therefore present results of analyses of pitch response transients and frequency response. In all cases, we compare the behavior of the nominal aircraft with the behavior of the damaged aircraft, after adaptation of the neural network.

5.3.1 Pitch response Transients

This transient analysis is important to determine the "overshoot" in pitch response. We compare the development of the pitch rate in the nominal and failure (with neural network active) case. For this simulation, off-nominal gains for the canards (in the range between –0.5 to 2) have been used. Each of the six responses in Figure 6, Figure 7, and Figure 8 compares a time history of pitch rate with the neural net on (red), versus neural net off (blue). The Figure 6 is for the maximum negative canard gain failure and the last is for the maximum positive gain failure. The active neural net reduces the overshoot much more at the negative end of the gain range than at the positive end. At the positive extreme, the neural net response essentially does not alter the system response. Thus, our analysis indicates that the neural net:

1. reduces the failure transient noticeably for the negative canard gains
2. does not reduce the smaller failure transients that result from positive canard gains
3. does not distinguishably alter the response of the aircraft to pilot input following the transient, i.e., its presence should not affect the handling qualities of the aircraft.

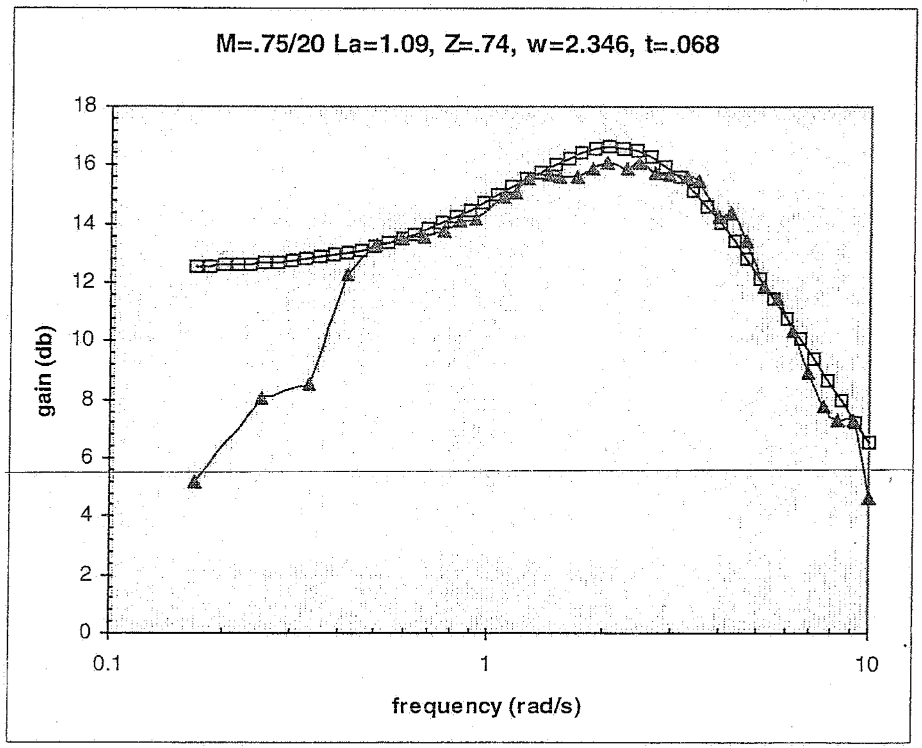

Figure 4. Equivalent System match of pitch response

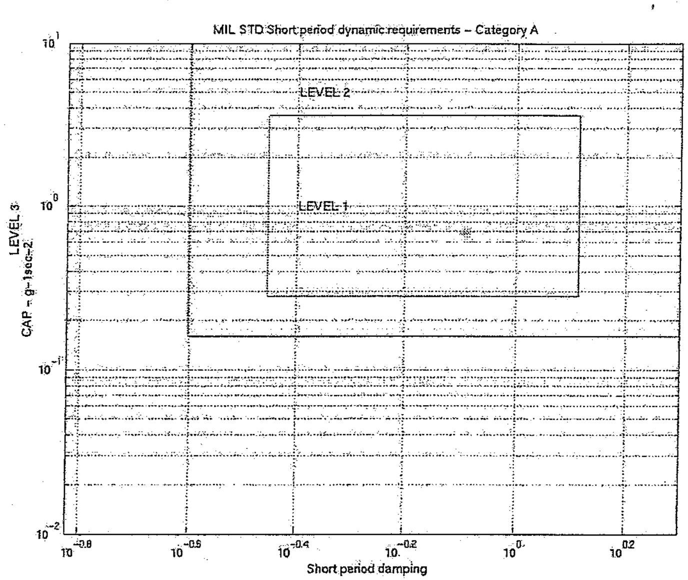

Figure 5. Control Anticipation Parameter (CAP)

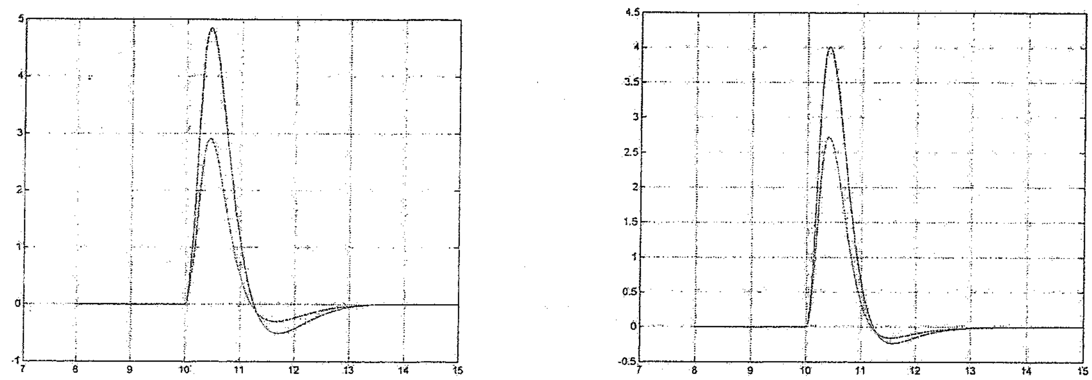

Figure 6 Pitch rate transients

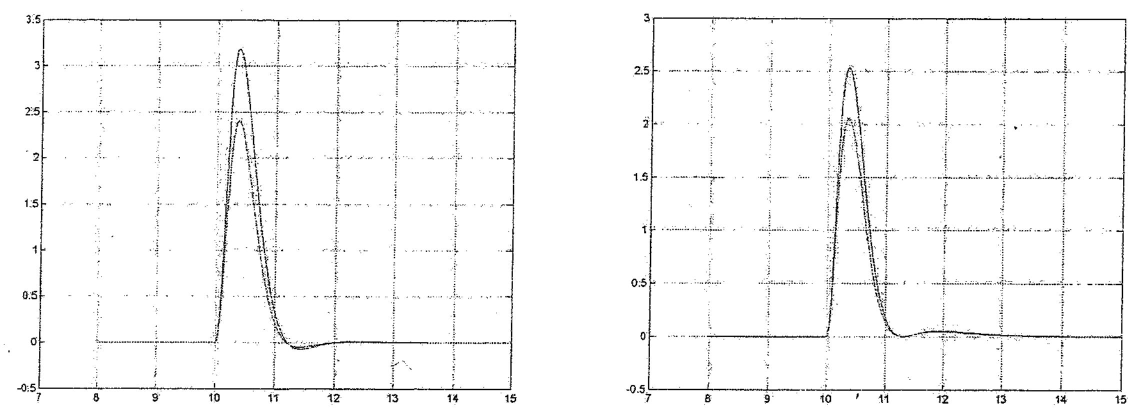

Figure 7 Pitch rate transients

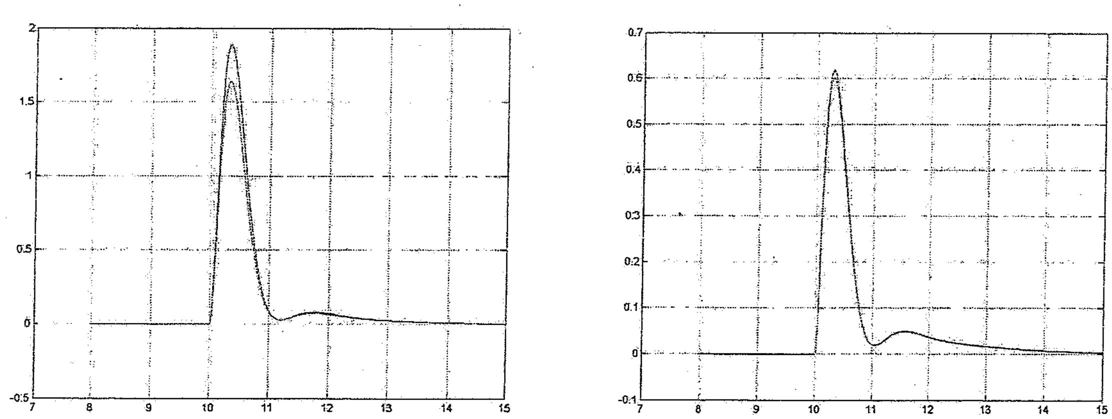

Figure 8 Pitch rate transients

A number of exploratory runs were made to determine how the neural net affected the response of pitch rate to piloted control, an important component of handling qualities. The results indicated that the neural net did not affect the response to control. This is illustrated in Figure 9. This figure incorporates the maximum negative canard gain, which as we have pointed out is the case where the net has the maximum effect on reducing the failure transient. Despite the net's strong effect in reducing the failure transient (red response), the doublet responses immediately following the failure indicate no difference due to the presence of the net.

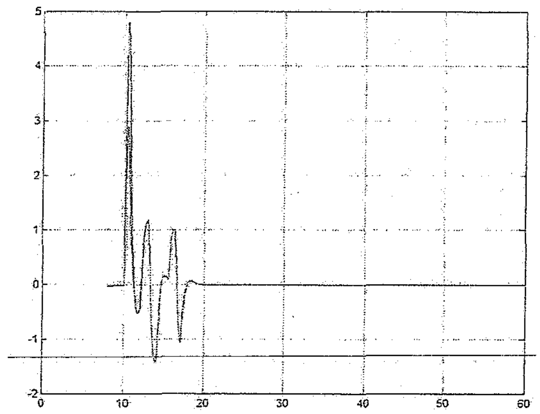

Figure 9 Negative canard gain failure and pilot doublets, net on (red) versus off (blue).

5.3.2 Frequency Response

In this experiment, we compare the frequency response of the LOE system with the full model. Figure 10 shows the response (measured in db, left) and the phase (in deg, right) over the frequency (in rad/s). Open squares indicate data, obtained by our LOE system, solid triangles correspond to our full model. In the nominal case, a good match between LOE and full model could be found.

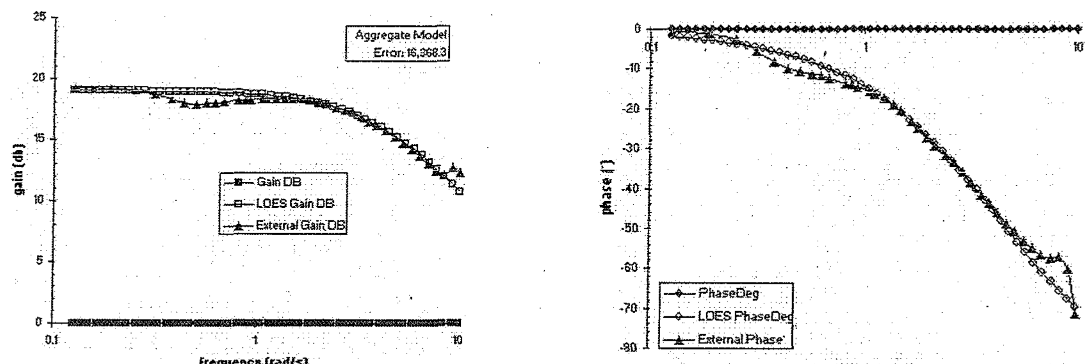

Figure 10. Equivalent system for nominal: Z= 1.1, Wn= 2.9 rad/sec, tau=.005 sec

Figure 10 shows a low-order-appearing response for the nominal aircraft with good damping (Z) and natural frequency (Wn) and time delay (tau). Figure 11 shows the situation for the damaged aircraft, after the network has been fully trained. The match between the LOES and the full model as well as the parameters are not as good as for the nominal case, but the high frequency lag does not decrease the natural frequency greatly and does not introduce a large delay. The aircraft is less low-order-appearing than the nominal, as shown by the worse match.

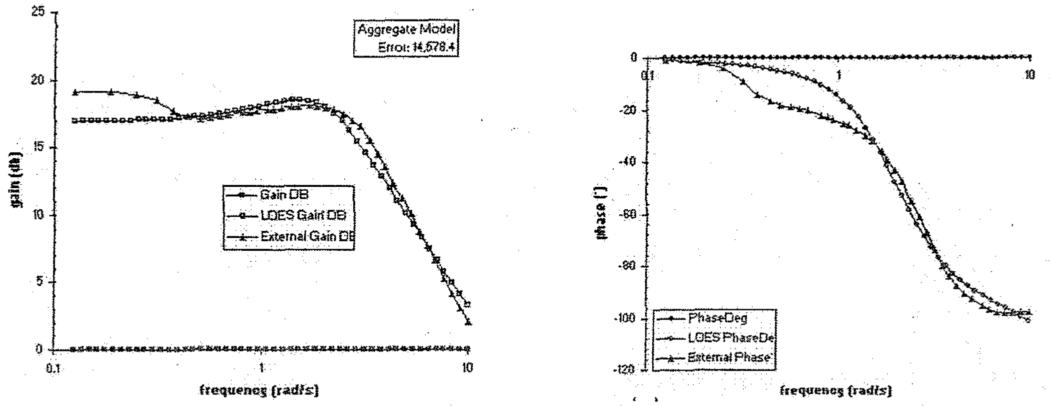

Figure 11. Equivalent system for damaged aircraft with fully trained neural network: Z=.63, Wn=2.02 rad/sec, tau=.026 sec

VI. Conclusions

In this paper, we have presented ongoing work on the performance analysis of neural network based adaptive controllers. We have presented simulation results of the Confidence Tool that dynamically calculates a performance metric for the neural network. Important aspects for handling quality, as laid out in MIL 1797 have been analyzed, using a low-order equivalent system of the aircraft and simulation results with the full model. Analysis of pitch response transients and frequency response have been discussed. For the nominal aircraft, a good match between the LOES and the full model could be established. In the failure case, the match is not as good, indicating that higher-order effects play a significant role in neuro-adaptive control—this aspect will be investigated in the future.

References

[1]Rysdyk, R. T., and Calise, A. J., "Fault Tolerant Flight Control via Adaptive Neural Network Augmentation," AIAA 98-4483, August 1998.

[2]Kaneshige, J., Bull, J., and Totah, J. J., "Generic Neural Flight Control and Autopilot System," AIAA-2000-4281, August 2000.

[3]Kaneshige, J. and Gundy –Burlet, K., "Integrated Neural Flight and Propulsion Control System," AIAA 2001-4386, August 2001.

[4]Cooper, G.E. and Harper, R. P., Jr. (1989)," The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities, NASA TN D-5153 "

[5] Hodgkinson, J, *Aircraft Handling Qualities*, Blackwell Science, 1999.

[6]Bishop, C.M., *Neural Networks for pattern Recognition*, Claderon Press, Oxford, 1995.

[7]Calise, A. and Rysdyk, R.T., " Nonlinear Adaptive flight control using neural networks, *IEEE Control Systems Magazine*, 21(6):14-26, 1998

[8]Mackall, D., Nelson, S., and Schumann, J., "Verification and Validation of Neural Networks for Aerospace Systems, " NASA/CR-2002-211409, July 2002.

[9]Gupta, P. and Schumann J. "A tool for verification and validation of neural network based adaptive controllers for high assurance systems," in *Proceedings IEEE High Assurance Software Engineering (HASE) 2004.*

[10]Schumann, J., Gupta, P., and Nelson, S., "On Verification and Validation of Neural Network Based Controllers," *Proceedings of Engineering Applications of Neural Networks (EANN'03)*, Malaga, Spain, September 2003.