

Operator Trust Function for Predicted Drone Arrival

Anna C. Trujillo¹

¹ NASA Langley Research Center,
MS 152
Hampton, VA, USA
anna.c.trujillo@nasa.gov

Abstract. To realize the full benefit from autonomy, systems will have to react to unknown events and uncertain dynamic environments. The resulting number of behaviors is essentially infinite; thus, the system is effectively non-deterministic but an operator needs to understand and trust the actions of the autonomous vehicles. This research began to tackle non-deterministic systems and trust by beginning to develop a user trust function based on intent information displayed and the prescribed bounds on allowable behaviors/actions of the non-deterministic system. Linear regression shows promise on being able to predict a person's confidence of the machine's prediction. Linear regression techniques indicated that subject characteristics, scenario difficulty, the experience with the system, and confidence earlier in the scenario account for approximately 60% of the variation in confidence ratings. This paper details the specifics of the linear regression model – essentially a trust function – for predicting a person's confidence.

Keywords: Trust function · Non-deterministic system · Linear regression · Autonomy · Confidence rating

1 Introduction

A primary goal for public and civil operators is to have one person managing several vehicles with different mission goals. To realize the full benefit from autonomy, these systems will have to react to unknown events and uncertain dynamic environments. The resulting number of behaviors is essentially infinite; thus, the system is effectively non-deterministic. So, rather than verify that an autonomous agent provides the correct answer in all cases, an impossibility with a non-deterministic system, instead determine whether verifying a defined solution space (i.e., bounds on system behavior) is feasible.

An operator overseeing a group of autonomous vehicles is a direct application of this approach. The operator needs to understand and trust the actions of autonomous vehicles. Achieving trust will become even more difficult and complicated if vehicles are able to make effectively non-deterministic decisions. An operator may learn to not trust or have confidence in such a vehicle if he is unable to understand why an autonomous vehicle is taking a particular action, which could result in system-wide failures and limited system performance due to the operator overriding the autonomous vehicle's protocols. Conversely, if the operator relies on autonomy in excess of its behavioral bounds, he may lose situation awareness of the system as a whole with consequences

ranging from suboptimal system performance to potentially catastrophic for conditions falling outside the autonomous vehicle's behavior boundary constraints. This research began to tackle these problems by trying to quantify the solution space non-deterministic systems inhabit as a function of the mission and then informing the operator of this solution space to foster trust and increase efficiency of the system as a whole.

The overall objectives were two-fold – (1) verifying bounds for non-deterministic decisions and (2) fostering trust in the actions taken by autonomous agents by making their decision process transparent to the operator. Objectives for this initial experiment were to (a) develop a non-deterministic system that operates within known bounds, (b) develop a display that shows possible outcomes from the current state, and (c) begin to develop a user trust function based on intent information displayed and the prescribed bounds on allowable behaviors/actions of the non-deterministic system. The autonomous agent had responsibility for mission performance that entailed trajectory planning and replanning in response to unanticipated events such as weather, other aircraft, etc. without outside operator supervision. The system behavior bounds were dependent on vehicle internal state self-knowledge, external environment represented by sensor data, and associated uncertainty. To have known and hence prescribed bounds, a modified Chua's circuit [1, 2] initially modeled the non-deterministic system. Autonomous agent intent information to display for user trust function initial development involved a drone arriving at its next waypoint at a specified latitude, longitude, altitude and time – a 4D trajectory.

2 Background

Research is just beginning on displaying possible outcomes and decision making (for example, see [3-5]) and previous research on trust focused on subjective measures (for example, see [6, 7]) rather than objective measures. Various models and questionnaires to define trust have been developed. For example, Hoff and Bashir [8] developed a three-layered framework for trust. This framework contains dispositional trust, which includes personality traits, situational trust, which encompasses workload, task difficulty and self-confidence, and learned trust, which includes experience with the system, knowledge of system performance and transparency. Others have found that personality traits may not significantly affect trust; instead, trust was more affected by the autonomous agent characteristics and the task characteristics [9].

The Army Research Laboratory has focused recently on agent transparency effects on operator trust [10-13]. They have found that increasing transparency increased operator's trust [10] but this increase is limited [11]. With this added information on agent transparency, increases in operator workload may occur due to additional information the operator must pay attention to; however, recent research indicated that added transparency information did not significantly increase workload [14, 15].

The Army Research Laboratory situation awareness-based agent transparency (SAT) model [12] mirrors Endsley's situation awareness model [16, 17] in that it has three levels. The first level consists of basic information such as purpose, process, and current performance and status. The second level consists of rationale or the agent's reasoning process which may include environmental and other constraints. The third level consists

of outcomes and includes projections of future outcomes, uncertainty, likelihood of success, and performance history.

Considering Chen's, et al. SAT model [12] and trust questionnaires [7-9], operator personality traits, autonomous agent history, and the task itself, this initial experiment looked to develop an objective trust function. This trust function could then be used to predetermine the needed information to provide to an operator to ensure trust is maintained among all agents and to maintain trust during an operation by changing the information provided to the operator based on signal variations. This will allow for an optimal system by ensuring that a lack of trust does not lead to an operator overriding the system [18] and that excessive trust does not lead to a lack of operator situation awareness [19].

This research looked to define objective measures that estimate user trust in a system. First, a non-deterministic system that operates within known bounds was designed and is described in sections 3.1 and 3.2. Second, a simple display indicating the predicted state of a drone arriving at a 4D gate was designed and is described in section 3.3. Lastly, a human-in-the-loop experiment, described in section 4, collected user trust of the system and this data was used to develop an initial trust function described in section 5.

3 Setup

3.1 Chua's Circuit

A Chua's circuit was used to mimic a non-deterministic system that operates within known bounds. Chua's circuits have values that are theoretically provable to fall within a defined range [1, 2, 20]. A basic Chua's circuit is shown in Fig. 1. The Chua diode consisted of resistors 1 thru 6. The gyrator, or inductor, consisted of resistors 7 thru 10 and the capacitor C . The values for these resistors, inductor, and capacitor were from [20]. C_1 , C_2 , and R depended on the scenario. Because four values were needed, the Chua's circuit was run twice with the same C_1 and C_2 values but with different R values. C_1 ranged from 8 nF (nanofarad) to 10 nF, C_2 ranged from 80 nF to 100 nF, and R ranged from 1800 Ohms to 2100 Ohms.

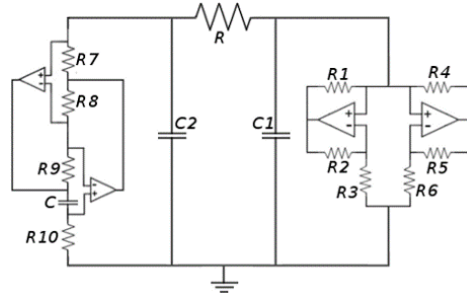
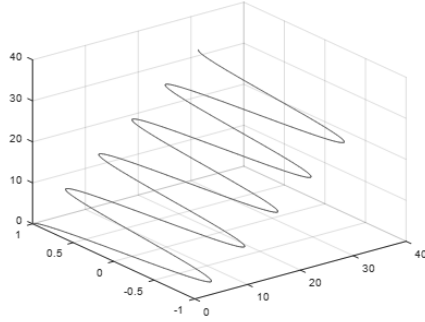


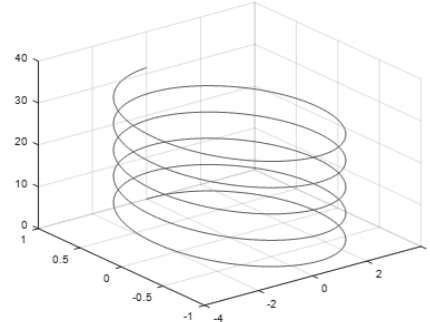
Fig. 1. Diagram of Chua's circuit (figure from <http://www.chuacircuits.com/matlabsim.php>).

3.2 Flight Paths

The drone followed four types of trajectory paths. The first was constant velocity latitude and longitude change with constant altitude. The second was constant velocity



(a) Boustrophedon flight path with constant changing altitude.



(b) Constantly ascending or descending helix flight path.

Fig. 2. Boustrophedon and helix flight paths.

latitude and longitude change with constant changing altitude (ascending or descending). The third consisted of a smooth boustrophedon pattern with constant changing altitude (ascending or descending) ((a)) and the last trajectory was a constantly ascending or descending helix ((b)).

The Chua's circuit simulated the drone's deviation from the flight path, specifically longitude, latitude, altitude, and time. The absolute maximum or bounds of the Chua's circuit to the prescribed flight paths had three variation levels – low, medium, and high (normalized latitude, longitude, altitude = ± 10 , ± 12 , ± 14 and time = ± 5 , ± 7 , ± 9 respectively) – and one scenario that actually exceeded the bounds at the end of the run – Exceed (latitude, altitude = ± 14 , longitude = ± 16 and time = ± 7). Fig. 3 depicts an example of the prescribed path and deviations from the prescribed path as generated from the Chua's circuit.

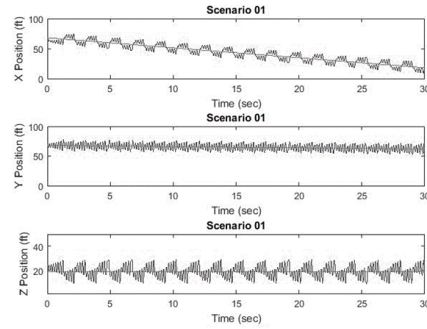


Fig. 3. Example constant velocity latitude and longitude change with constant altitude flight path with low variation.

3.3 Displays

Each subject saw several different displays – (1) trajectory display only, (2) trajectory display plus predicted deviation from prescribed gate, (3) previous plus deviation bars, (4) previous plus predicted deviation values from prescribed gate, and (5) previous plus likelihood of predicted deviation values from prescribed gate.

Trajectory Display. The trajectory display showed the current position of the drone relative to its prescribed latitude, longitude, and altitude positions (left side of Fig. 4).

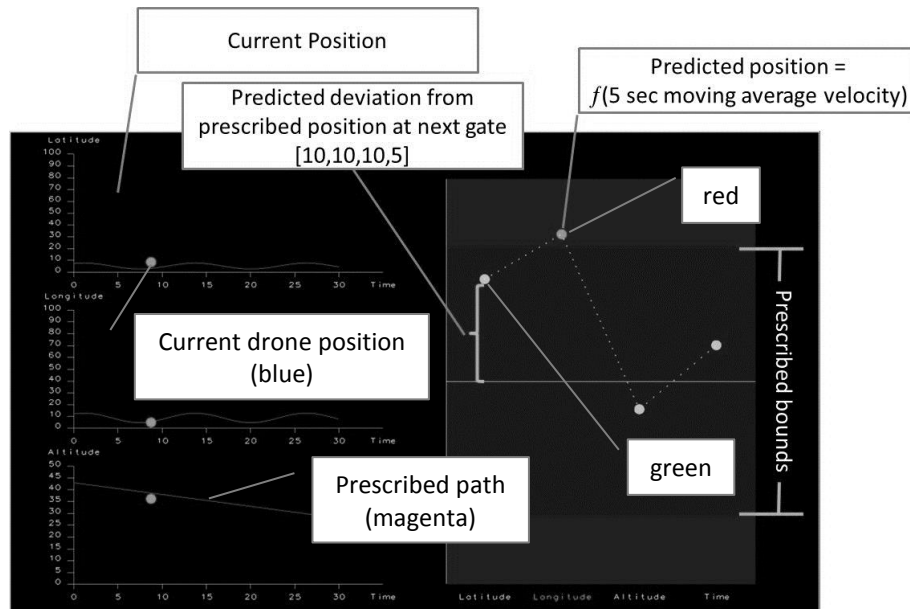


Fig. 4. Trajectory display plus predicted deviation from prescribed gate. Predicted position deviation is normalized. Prescribed bounds area is a green shade while outside the prescribed bounds area is a red shade. Predicted position dot and its text is same color as the area it is in.

Trajectory Display plus Predicted Deviation from Prescribed Gate. This display added a panel to show the weighted average predicted deviation from the prescribed gate which was to be reached at the end of the run (right side of Fig. 4). The green area indicated the acceptable prescribed bounds which were ± 10 units for latitude, longitude and altitude and ± 5 sec for time. The red areas were exceedances of these bounds. The predicted position was a function of the 5 sec moving average velocity of latitude, longitude, altitude, and time. If a value was predicted to exceed the bounds when the drone was to reach the gate, the dot and its associated text were colored red otherwise they were green. The light dotted white line between the values was there to aid the subject in lining up the dot location with its heading text.

Trajectory Display plus Predicted Deviation from Prescribed Gate with Deviation Bars. This display added deviation bars that indicated the range of possible values (Fig. 5). The deviations were a function of average velocity and velocity of the 4D gate. The high and low values did not have to be equal.

Trajectory Display plus Predicted Deviation from Prescribed Gate with Deviation Bars and Predicted Deviation Values. This display added predicted deviation values for the predicted value and its high and low value (Fig. 5). As with the dots being color coded, the values were the same color as where the value resided. So, if the value was within the prescribed bounds, the number was green; otherwise the number was red.

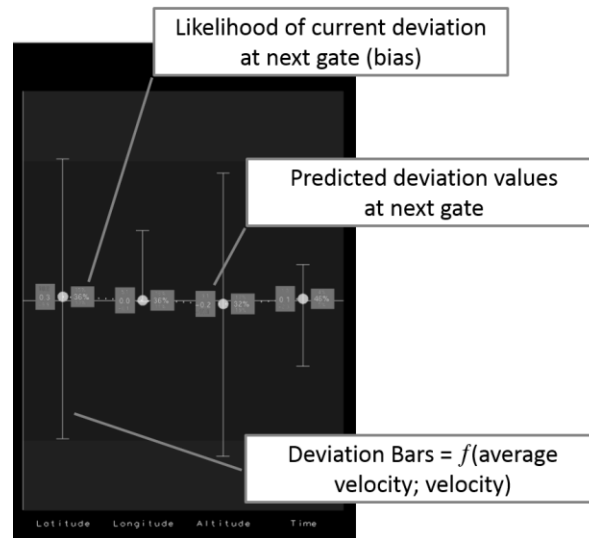


Fig. 5. Predicted deviation from prescribed gate with deviation bars, predicted deviation values at next gate, and likelihood of current deviation at next gate. As with Fig. 4, text color matches the area the value is in.

Trajectory Display plus Predicted Deviation from Prescribed Gate with Deviation Bars, Predicted Deviation and Likelihood of Predicted Deviation Values. The final display added the likelihood of the predicted deviation values (Fig. 5). As before, the percentages were the same color as the predicted deviation values.

3.4 Confidence Rating Question

At three points during each data run (at 10 sec, 20 sec, and at the end of the run), the scenario paused so that subjects could answer “How confident are you that the drone would reach its gate?” This question was on a scale of 0, no confidence that the drone would reach its gate within the prescribed boundaries, and 100, absolutely sure that the drone would reach its gate within the prescribed boundaries.

3.5 Electronic NASA-TLX Questionnaire

At the end of each run, subjects completed an electronic version of the NASA Task Load Index (NASA-TLX) [21]. This questionnaire was always the last set of questions presented at the end of the 30 sec run.

4 Procedure

Each subject had 15 data runs – 3 with each of the displays described in section 3.3. Each data run lasted for 30 sec and at the 10 sec, 20 sec, and 30 sec points, the scenario

paused so that the subject could answer the questions described 3.4. At the end of the run, subjects also completed an electronic version of the NASA-TLX. At a change of display, subjects had 2 practice runs that behaved just like the data runs.

5 Results for End Confidence Rating

Data was analyzed using IBM® SPSS®¹ V24 automatic linear regression techniques. Significance was taken at $p \leq 0.05$.

The linear regression to predict confidence rating at the end of the run had an accuracy of 62%. The significant factors are earlier confidence ratings during the run, subject personality, path deviation level, when the run occurred, and an intercept of 36 (Eq. 1).

$$CR_{end} = 36 + 0.24CR_{20sec} + 0.21CR_{10sec} + subject_{personality} + path_{deviation} + run_n \quad (1)$$

where

$$\begin{aligned} CR &= \text{confidence rating at end, 20 sec, or 10 sec} \\ subject_{personality} &= \begin{cases} 15 & \text{if low frustration and good performance} \\ 9 & \text{if mid frustration and mid performance} \\ 0 & \text{if high frustration and poor performance} \end{cases} \\ path_{deviation} &= \begin{cases} 0 & \text{if deviation} \neq \text{high} \\ -9 & \text{if deviation} = \text{high} \end{cases} \\ run_n &= \begin{cases} 6 & \text{if } n \text{ is early or late} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

5.1 Earlier Confidence Ratings Effects on End Confidence Rating

From Eq. 1, the confidence rating at the 10 sec interval affected the end confidence rating by a factor of 0.21 ($p \leq 0.01$; importance = 0.15) and the confidence rating at the 20 sec interval affected the end confidence rating by a factor of 0.24 ($p \leq 0.01$; importance = 0.16). As can be seen in Fig. 6, the confidence ratings during a run increase as the run continues. Thus, as the run continues, newer information influences the end confidence rating the most.

¹The use of trademarks or names of manufacturers in this report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.

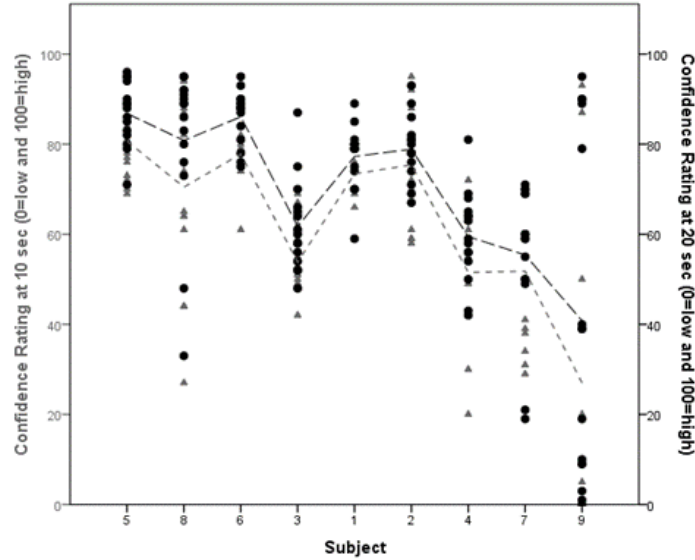


Fig. 6. Confidence rating at 10 s into run (gray triangles) and at 20 s into run (black circles) by subject. The gray dotted line indicates the average confidence rating at 10 sec and the black dotted line indicates the average confidence rating at 20 sec.

5.2 Subject Effects on End Confidence Rating

Subject personality had a significance of $p \leq 0.01$ and an importance of 0.30. There were 3 groups of subjects (Fig. 7). In general, subjects that had low frustration and good performance added 15 points to their end confidence rating and subjects that had some frustration and slightly lower performance with the task added 9 points to their end confidence rating. Subjects that had high frustration and poor performance on the task added no additional points to their end confidence rating.

5.3 Scenario Level Effects on End Confidence Rating

Scenario level, or path deviation level, had a significance of $p \leq 0.01$ and an importance of 0.24 (Fig. 9). Scenarios which had a path deviation of at least ± 14 in latitude, longitude, and altitude and ± 9 in time decreased the end confidence rating by nine. Not surprisingly, vehicles that deviate quite a bit from the planned path resulted in a lower confidence rating.

5.4 Run Occurrence Effects on End Confidence Rating

Finally, when the run occurred had a significance of $p \leq 0.02$ and an importance of 0.10 (Fig. 9). The first and last runs, which added 6 points to the end confidence rating, were in a separate group from the middle runs. This may indicate an operator attentional lag in the middle of a mission.

5.5 Linear Regression End Confidence Rating Prediction

Fig. 10 shows an example predicted end confidence rating for a subject in group 1 by run number. Absolute error was calculated using Eq. 2. The mean absolute error was 5.5 with a standard error of the mean of 1.6, maximum absolute error of 24.3 and a minimum absolute error of 0.2. Table 1 indicates the above values for all subjects. As can be seen in Table 1 and Fig. 10, the absolute mean error is approximately 10 indicating that the linear regression equation can predict the end confidence rating of a person fairly accurately.

$$Absolute\ Error = \left| \frac{End\ Confidence\ Rating_{[Subject, Run]} - Predicted\ End\ Confidence\ Rating_{[Subject\ Personality, Run_{Group}, Path_{Deviation}]}}{1} \right| \quad (2)$$

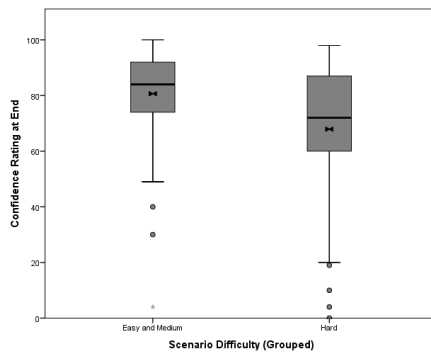


Fig. 9. End confidence rating by grouped scenario difficulty box plot. Bow tie indicates the mean.

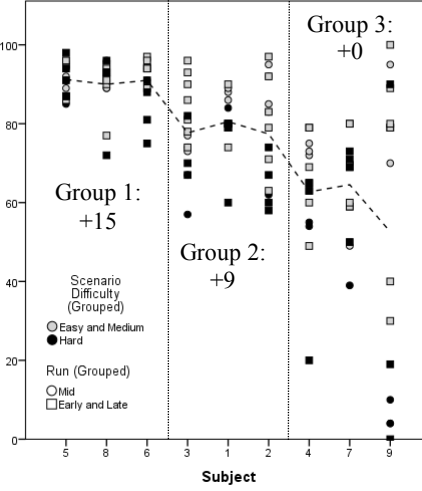


Fig. 7. End confidence rating by subject. The black dotted line indicates the average end confidence rating for each subject. Light grey indicates easy and medium scenario difficulty and black indicates hard scenario difficulty. Circles indicate mid-runs and squares indicate early and late runs.

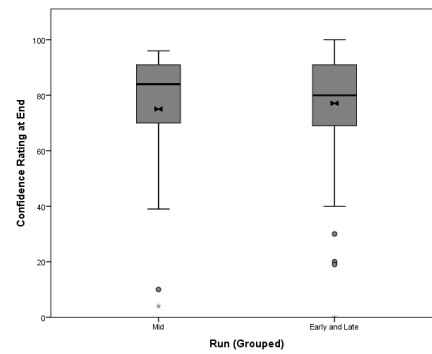


Fig. 9. End confidence rating by grouped run occurrence box plot. Bow tie indicates the mean.

Table 1. End confidence rating mean error statistics for each subject. Absolute error is calculated using Eq. 2.

Subject	Mean Absolute Error	Absolute Standard Error of the Mean	Absolute Maximum Error	Absolute Minimum Error
5	5.5	1.6	24.3	0.2
8	9.5	2.8	40.4	0.5
6	6.0	0.9	13.8	1.5
3	9.8	1.8	23.0	1.2
1	7.0	1.4	17.4	0.1
2	8.7	1.5	20.7	0.0
4	8.3	1.5	21.3	0.2
7	10.4	1.8	22.3	0.3
9	24.6	3.6	43.1	0.1

6 Discussion

The above results indicate that an operator's confidence or trust can be predicted by objective measures (see Eq. 1). Each successive confidence rating builds on previous confidence ratings for a particular run. The function also is dependent on the subject's personality with regards to workload, how easily he may become frustrated, and his performance. The time or experience within a mission also affects trust. Lastly, if the vehicle has large deviations from the proscribed path, confidence decreases. Although the information provided on the display did not affect the linear regression, it did highlight the deviations; therefore, this variable may have rolled up into the scenario deviation variable. In general,

$$Trust = f(personality, observed\ path\ deviation, experience).$$

7 Conclusions

A primary goal for public and civil operators is to have one person managing several vehicles with different mission goals. To realize the full benefit from autonomy, these

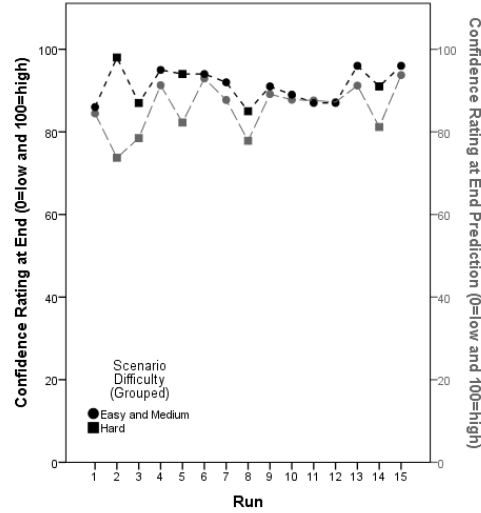


Fig. 10. Example end confidence rating of a subject (black icons) compared to predicted end confidence ratings (gray icons). Circles indicate easy and medium scenario difficulty. Squares indicate hard scenario difficulty.

systems will have to react to unknown events and uncertain dynamic environments. The resulting number of behaviors is essentially infinite; thus, the system is effectively non-deterministic.

Even with the system effectively non-deterministic, an operator needs to understand and trust the actions of the autonomous vehicles in order for the system as a whole to operate optimally. This research began to tackle non-deterministic systems and trust by beginning to develop a user trust function based on intent information displayed and prescribed bounds on allowable behaviors/actions of the non-deterministic system. Linear regression shows promise on being able to predict a person's confidence of the machine's prediction. Linear regression techniques indicated that subject characteristics, scenario difficulty, the experience with the system, and confidence earlier in the scenario account for approximately 60% of the variation in confidence ratings.

However, these results are preliminary because this experiment entailed a few short runs with a limited subject pool. Additional and longer runs will better determine how time affects confidence. A larger subject pool will aid in determining more precise subject characteristics that affect confidence. This will simplify the trust function in that it will not have to be tuned to each individual. Furthermore, even though varying amounts of information were provided, the primary driving factors appears to be time and the amount of deviation from path. The information provided on the displays may have been rolled into the deviation from path variable. Therefore, additional research needs to be conducted in order to better refine the effects of deviation and the information provided.

In any case, trust appears to be a function of personality, deviation, and time. With a parametric function, user trust could be predicted. With this prediction, additional information could be provided to maintain an appropriate level of trust. The appropriate level of trust maintained among team members, whether they be human or machine, will help enable a system to perform optimally. Lastly, providing detailed information regarding the reasoning behind the prediction (second level of the SAT model) may also beneficially affect the trust function.

Acknowledgments. This research was supported by NASA Langley Research Center IRAD funding in 2016.

References

1. Chua, L.O.: Chua Circuit. *Scholarpedia* 2, 1488 (2007)
2. Chua, L.O., Wu, C.W., Huang, A., Zhong, G.-Q.: A Universal Circuit for Studying and Generating Chaos-Part I: Routes to Chaos. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications* 40, 13 (1993)
3. Beller, J., Heesen, M., Vollrath, M.: Improving the Driver–Automation Interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 55, 11 (2013)
4. McGuirl, J.M., Sarter, N.B.: Supporting Trust Calibration and the Effective Use of Decision Aids by Presenting Dynamic System Confidence Information. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 48, 10 (2006)
5. Verberne, F.M.F., Ham, J., Midden, C.J.H.: Trust in Smart Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 54, 11 (2012)

6. Couch, L.L., Jones, W.H.: Measuring Levels of Trust. *Journal of Research in Personality* 31, 18 (1997)
7. Jian, J.-Y., Bizantz, A.M., Drury, C.G.: Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics* 4, 16 (2000)
8. Hoff, K.A., Bashir, M.: Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 407-434 (2015)
9. Schaeffer, K.E.: The Perception and Measurement of Human-Robot Trust. Department of Modeling and Simulation in the College of Sciences, vol. Doctor of Philosophy, pp. 359. University of Central Florida, Orlando, Florida (2013)
10. Boyce, M.W., Chen, J.Y.C., Selkowitz, A.R., Lakmani, S.G.: Effects of Agent Transparency on Operator Trust. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, pp. 179-180. ACM, New York, NY, USA (2015)
11. Chen, J.Y.C., Barnes, M.J., Selkowitz, A.R., Stowers, K.: Effects of Agent Transparency on Human-Autonomy Teaming Effectiveness. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC) pp. 1838-1843. IEEE, (Year)
12. Chen, J.Y.C., Procci, K., Boyce, M.W., Wright, J., Garcia, A., Barnes, M.J.: Situation Awareness-Based Agent Transparency. In: Laboratory, U.S.A.R. (ed.), pp. 36. U.S. Army Research Laboratory, Aberdeen Proving Ground, MD (2014)
13. Lakhmani, S., Abich, J., Barber, D., Chen, J.: A Proposed Approach for Determining the Influence of Multimodal Robot-of-Human Transparency Information on Human-Agent Teams. In: Schmorow, D.D., Fidopiastis, C.M. (eds.) *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience: 10th International Conference, AC 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings, Part II*, pp. 296-307. Springer International Publishing, Cham (2016)
14. Mercado, J.E., Rupp, M.A., Chen, J.Y.C., Barnes, M.J., Procci, K.: Intelligent Agent Transparency in Human-Agent Teaming for Multi-UxV Management. *Human Factors* 58, 401-415 (2016)
15. Wright, J.L., Chen, J.Y.C., Barnes, M.J., Hancock, P.A.: Agent Reasoning Transparency's Effect on Operator Workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60, 249-253 (2016)
16. Endsley, M.R.: Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors* 37, 32-64 (1995)
17. Endsley, M.R.: Situation Awareness Misconceptions and Misunderstandings. *Journal of Cognitive Engineering and Decision Making* 9, 4-32 (2015)
18. Parasuraman, R., Riley, V.: Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39, 230-253 (1997)
19. Moray, N., Inagaki, T., Makoto, I.: Adaptive Automation, Trust, and Self-confidence in Fault Management of Time-critical Tasks. *Journal of Experimental Psychology: Applied* 6, 44-58 (2000)
20. <http://www.chuacircuits.com/>
21. Trujillo, A.C.: How Electronic Questionnaire Formats Affect Scaled Responses. 2009 (15th International Symposium on Aviation Psychology, Dayton, OH (2009))