EXAMINING CONTROLLER TASK REDISTRIBUTION AS AN INDICATOR OF PEAK WORKLOAD

Paul U. Lee and Joey Mercer
SJSU / NASA Ames Research Center
Moffett Field, CA, 94035-1000
{paul.u.lee, joey.mercer}@nasa.gov

Controller workload is a key factor in limiting en route capacity that has been modeled through fast-time modeling, real-time simulations, and operational data. In most of these efforts, the focus has been on correlating workload with other "objective" metrics, such as number of clearances, number of aircraft, etc. A missing component from such analyses is the controller's strategy for managing workload. Workload is not a passive factor that mirrors the controller actions (e.g. handoffs, clearances). Instead, controllers actively moderate and re-distribute the types and the frequency of actions based on their perceived workload. In this paper, we examine this strategy shift by associating bookkeeping tasks and route/altitude clearances with online workload ratings. Overall, the data suggest that only in high traffic scenarios in which the controller workload approached the maximum threshold,, the controllers shed peripheral tasks related to monitoring and bookkeeping as the traffic increases and their perceived workload transitions from low to high. Whenever workload reached a maximum, some bookkeeping tasks were delayed and performed in "groups" after the peak traffic subsided.

INTRODUCTION

In developing the future Next Generation Air Transportation System (NextGen), controller workload has been identified as a key limiting factor for the significant increase in capacity that has been predicted, and therefore has been an active field of research (e.g. Athenes, Averty, Puechmorel, Delahaye, & Collet, 2002; Hilburn, Bakker, Pekela, & Parasuraman, 1997; Stein, 1985). Because workload ratings are subjective and highly prone to individual differences, some researchers have tried to replace workload with more objective metrics, such as aircraft count, number of altitude changes, number of coordination events, traffic geometry, total time in sector, etc. In related research the collective effect of all factors that contribute to air traffic control complexity was examined and termed "dynamic density" (Kopardekar & Magyarits, 2003). One of the key motivations for dynamic density research is to find a set of metrics that can replace current day Monitor Alert Parameters (MAPs) to predict traffic complexity and associated controller workload.

Significant efforts in correlating workload with objective task load metrics miss a key causal relationship between workload and those task metrics: controllers often monitor their own workload to actively redistribute their task priorities, which in turn helps to manage and maintain their workload at an acceptable level. For example, controllers may try to provide efficient routings during low traffic situations, even if it requires more workload but abandon such practices during high traffic/workload situations in favor of actions that minimize workload.

The idea that workload is something that controllers actively manage has been noted by Sperandio (1978). He suggested that controllers kept their workload within acceptable limits by changing their task/resource management strategies. He examined how air traffic controllers modified their cognitive processes when the number of aircraft controlled increased. As the traffic level increased, controllers progressively used procedures that had less cognitive cost in order to minimize the workload impact. In addition, controllers sacrificed secondary objectives in high traffic density in order to maintain separation and other principal objectives.

Over-the-shoulder observations from a number of recent human-in-the-loop simulation studies have indicated a similar pattern of controllers' workload management strategies. In general, controllers seemed to re-distribute controller tasks as traffic level increased. For example, handoff-related tasks, which are closely related to aircraft count, increased linearly as the traffic increased. In contrast, controllers tended to shed bookkeeping and monitoring tasks during peak workload while still trying to maintain efficient traffic flow management. If the traffic reached an unmanageable level even after shedding these tasks, the controllers went into a "survival mode" in which they solely focused on maintaining adequate safe separation. By effectively shedding lower priority tasks as the traffic increased, the controllers seemed to effectively keep their workload below the maximum threshold.

To examine the moderating effect of workload on the types and number of tasks a controller is able to perform, we re-analyzed the data from a prior study. In this study, high traffic scenarios were used to evaluate En Route Free Maneuvering concept element in NASA's Distributed Air-Ground Traffic Management (DAG-TM) project (Lee, Prevot, Mercer, Smith, & Palmer, 2005). DAG-TM studies were conducted with an assumption of far-term equipage levels, including fully integrated advanced air and ground decision support tools (DSTs) with data link. Transfer of communication was also automated and integrated with data link, sending the frequency change uplink message to the flight deck with the handoff acceptance of the next sector.

A time-series plot was used to compare the task load with the associated workload ratings. The study, analyses, and the results are described below.

METHOD

Participants

The experiment included 22 commercial airline pilots and 5 certified professional air traffic controllers. Four controllers staffed four radar positions (three high altitude sectors and one low altitude sector). One additional controller served as a tracker/supervisor to support the radar controllers during peak workload periods. Twenty-one aircraft simulators were flown by participant pilots at NASA Ames and NASA Langley. All remaining aircraft in the simulation were flown by pseudo-pilots with autonomous agent support at NASA Ames and NASA Langley.

Airspace

The simulation airspace included portions of Albuquerque Center (ZAB), Kansas City Center (ZKC), Fort Worth Center (ZFW) and Dallas-Fort Worth TRACON (DFW) (Figure 1). Controller participants worked four test sectors in the northwest arrival corridor: three high altitude sectors (Amarillo in ZAB, Wichita Falls and Ardmore in ZFW), and one ZFW low altitude sector (Bowie). Three retired controllers handled the surrounding traffic that entered or exited the test sectors.

The overall scenario traffic was composed of a mixture of arrivals, departures and overflights. Arrivals transitioned from the northwest (Amarillo high and Wichita Falls High) and from the north (Ardmore high). The two main streams of arrivals merged at the BAMBE meter fix in the Bowie low sector before entering the TRACON. The traffic mix in Amarillo consisted of arrivals and overflights in level flight. Wichita Falls traffic was mostly arrivals while Ardmore had arrivals, departures, as well as a significant number of overflights.

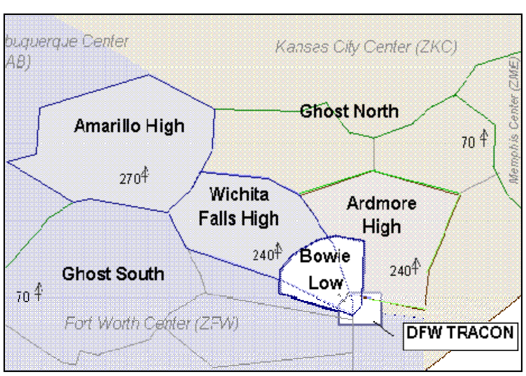

Figure 1. Simulated airspace

Ground Capabilities

To maximize the benefits of the advanced air and ground-side decision support tools (DSTs), they were integrated with Controller Pilot Data Link Communication (CPDLC) and the Flight Management System (FMS). This integration allowed controllers and pilots to exchange 4-D trajectory information quickly and with a reduced impact on workload relative to the conventional means of exchanging this information via voice. The controller decision support tools have been integrated into a high fidelity emulation of the Display System Replacement (DSR) controller workstation. This DSR emulator is highly configurable and scalable with the ability to mimic both DSR workstations in the field today and future DSRs with advanced DSTs. In order to support the concept, all aircraft were equipped with CPDLC, FMS, and automatic dependent surveillance-broadcast (ADS-B). The aircraft flown by the commercial pilot participants also had Cockpit Display of Traffic Information (CDTI) displays integrated with conflict detection & resolution (CD&R) and advanced required time of arrival (RTA) capabilities. More detailed descriptions of the ground capabilities are presented in Lee, Prevot, Mercer, Smith, & Palmer (2005).

Experimental Conditions

There were four experimental conditions in this study, two of which are re-analyzed and discussed in this paper. In the first “high traffic” condition, the aircraft count reached a peak that was higher than current day MAP values resulting in a peak workload level that was beyond the allowable level in current day operations. In the second “moderate traffic” condition, the peak aircraft count was reduced by approximately five aircraft, resulting in a comfortable moderate workload for the controller participants. Four data collection runs were analyzed per condition.

Subjective workload assessments were collected from controllers with the Workload Assessment Keyboard using the Air Traffic Workload Input Technique, or ATWIT (Stein, 1985). Controllers were required to rate their workload on a scale of 1 to 7 at 5-minute intervals throughout each simulation run. In the same five minute intervals, various task load metrics such as clearances, handoffs, and average aircraft count during the interval were tabulated. The results are described in the following section.

RESULTS

Workload vs. Aircraft Count. Figure 2 illustrates the traffic pattern for Amarillo high sector during the high and moderate traffic conditions. The graph shows that the peak aircraft count (averaged across 5-minute time span) reached 23 and 17 aircraft respectively for the high and the moderate traffic conditions. The results showed that a difference of five aircraft between the two traffic scenarios resulted in large differences in workload ratings. High traffic scenarios resulted in average workload ratings between 4 and 5 which correspond to high workload for most controllers, as they generally reserve 6 and 7 ratings to report situations with catastrophic failures or major re-planning (e.g. heavy thunderstorm or airport closures). In contrast, moderate traffic scenarios resulted in workload ratings between 2 and 3 which correspond to low to moderate workload.

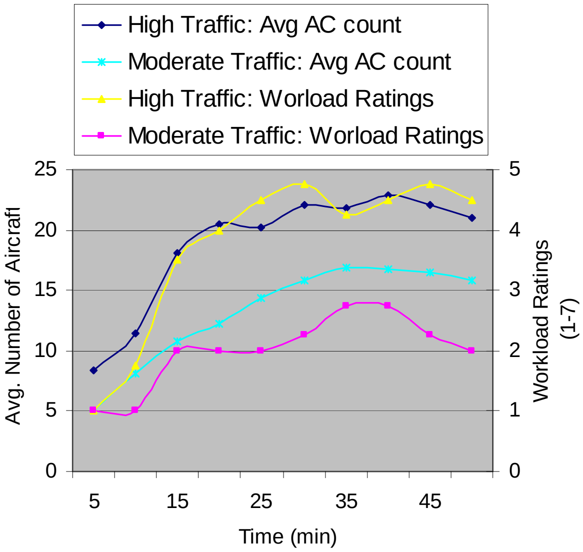

Figure 2. Number of aircraft count vs. workload ratings for high and moderate traffic scenarios

Workload vs. Monitoring/Bookkeeping Tasks. The tasks were divided into handoff-related, clearances, and monitoring/bookkeeping tasks. In this section, we focus on the monitoring/ bookkeeping tasks because previous findings (e.g. Sperandio, 1978) suggest that controllers shed secondary/less essential tasks as traffic increases.

From the monitoring/bookkeeping task data that was available from each run, only the following occurred with enough frequency to be analyzed: FMS route display, datablock adjustment, and datablock toggle. These tasks correlate strongly with one or more bookkeeping tasks that controllers engage in to maintain situation awareness. For example, FMS routes were displayed when controllers wanted to see their routes to resolve potential conflicts or to provide service to the aircraft.

Figures 3a and 3b suggest controllers display routes during the earlier phases of the scenarios when they have more time to plan their actions strategically. In Figure 3a, the data in the high traffic scenarios suggest that controllers display the FMS routes for a peak of 20% of the aircraft 15 minutes into the scenario and less so thereafter as their perceived workload increased. Fig. 3b shows a similar result for the moderate traffic scenarios in which controllers displayed FMS routes at a peak (approx. 15% of the aircraft) at 15 minutes into the scenario. Both data show a slight increase in the task frequency during dips in the workload ratings suggesting that this task is done when workload permits.

One of the interesting results from the analysis of workload and its relationship to task performance was the frequency of datablock adjustments in the high traffic scenarios. One reason for datablock adjustments was to organize the datablocks so that a plane entering or exiting a sector had a particular datablock orientation to remind the controllers of its current status. The other reason for the adjustments was to minimize display clutter by keeping the datablocks from overlapping on the screen. For these reasons, one would predict that datablock adjustment per aircraft would either stay relatively constant or increase slightly with increased levels of traffic. However in the *high traffic scenarios*, the data suggest that the frequency of this task sharply decreases (from about 60% of the aircraft to 10%) as the workload increases to its peak (around 30 minutes into the scenario) and increases again as the scenario continues (see Figure 3a). The results support the hypothesis that controllers shed this task when the workload is high, presumably because it is a lower priority task during peak workload. What is interesting, however, is that this pattern of results is not duplicated in the *moderate traffic scenarios*. Figure 3b shows that although the frequency of datablock adjustment per aircraft dropped to approximately 20% at 15 minutes, the rate climbed back up between 30% and 50% for the rest of the scenario. Combined results suggest that high (but not moderate) workload situations reduce the rate of datablock adjustments, likely because controllers manage their workload by minimizing the frequency of peripheral tasks.

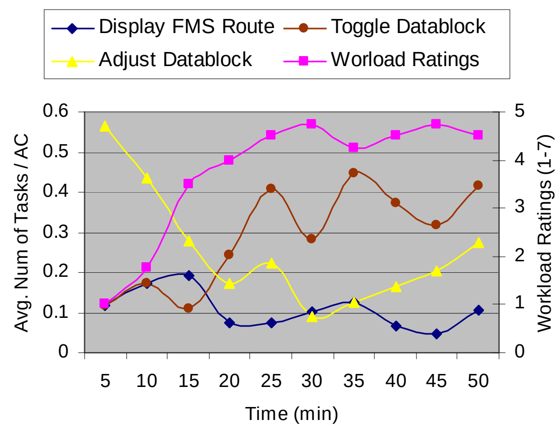

Figure 3a. Number of "bookkeeping" tasks vs. workload ratings for high traffic scenarios

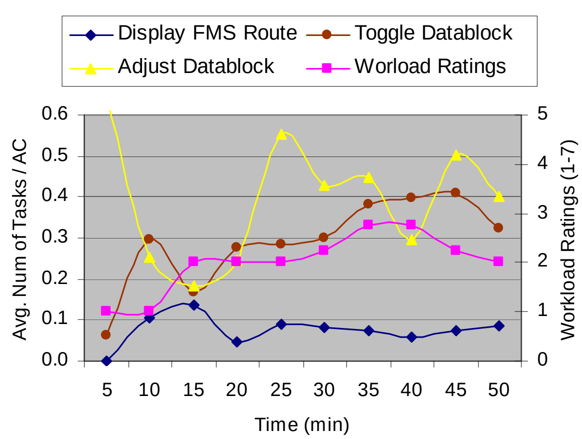

Figure 3b. Number of "bookkeeping" tasks vs. workload ratings for moderate traffic scenarios

Figure 4 highlights the difference in the datablock adjustment patterns in high vs. moderate traffic levels. Data are aggregated across three time periods. During the first ten minutes, aircraft count and workload are both low (see Figure 2). During the next fifteen minutes (10-25min of the scenario), both aircraft count and the associated workload transition from low to the peak count/workload for that particular traffic scenario. Finally, the aircraft count and workload maintain their peak during the subsequent twenty minutes (25-45min).

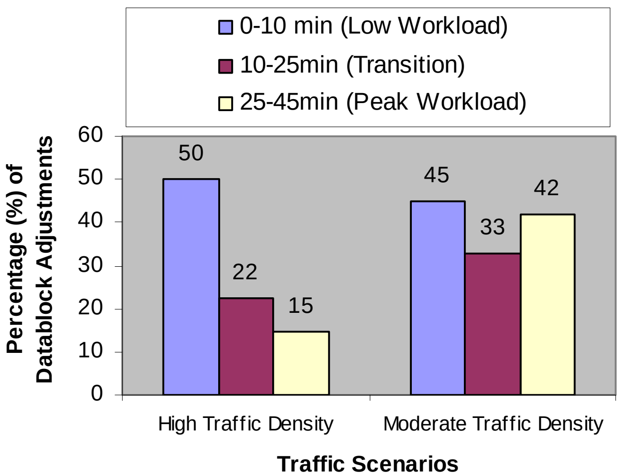

Figure 4. Percentage of datablock adjustments at different traffic/workload levels for high and moderate traffic density

By examining the percentage of datablock adjustments by these different stages of traffic and workload patterns, one can see a clear difference of datablock adjustment patterns between high and moderate traffic density. Figure 4 shows that in high traffic density, controllers adjusted datablocks on 50% of the aircraft at the beginning when there were few aircraft, but the percentage dropped to 22% as the traffic ramped up and dropped further to 15% at the peak traffic level. In contrast, controllers in moderate traffic density reduced the datablock adjustments from 45% to 33% as the traffic initially ramped up, but the percentage of adjustments returned to 42% once the traffic/workload reached its peak. The data suggest that in moderate traffic conditions, controllers had enough mental resources to manage their workload without sacrificing datablock adjustments (i.e. a secondary task).

The datablock toggle task is another secondary task that shows an interesting finding. In this simulation environment, toggling a datablock almost exclusively had one function – to minimize the datablock after the plane has been handed off to the downstream sector. This was an important task to perform as it minimized display clutter, but the timing of the event seemed to be less critical. Given this understanding, it was interesting to see an “oscillation” pattern in this task that was out of phase with workload data in the *high traffic scenarios*. The result suggests that whenever the workload was at its peak, controllers delayed minimizing the datablock until the workload dipped slightly from its peak (see Figure 3a). In contrast, the oscillating pattern was not present in the *moderate traffic* condition, suggesting that this task was not delayed when the workload was not at its maximum.

Workload vs. Route/Altitude Clearances. The data used for the analysis were collected for a future operational environment that allowed route and altitude trial planning capabilities that could construct conflict-free 4-D trajectories graphically using the trackball. Controllers could then data link the conflict-free paths to the flight deck bypassing voice clearances altogether. The only situations that required voice commands were when the clearances needed to be delivered right away or if the pilot had not responded to data link clearances. In such cases, controllers verbally assigned regular and interim/temporary altitudes as well as heading and speed changes.

If the controllers minimize their workload in high traffic/workload situations by shedding lower priority and higher workload tasks, they would likely choose altitude over lateral clearance during high workload situations since altitude clearances require less workload than a lateral route clearance. When the workload is moderate, our controller participants have commented that they would generally try a lateral solution first as they try to leave the planes at their current altitudes. They also wanted to have the altitude solution available as an “out” maneuver if any last-minute maneuvers are needed. Given these two constraints, one would expect that there would be a greater number of lateral route maneuvers during the low to moderate workload situations and a shift to a greater number of altitude maneuvers during high traffic situations.

The results from the altitude and lateral route clearance generally support the above hypothesis but the details of the data are difficult to interpret. In general, the data suggest that the route clearances are used more often during low to moderate workload situations and the altitude clearances are used most often during peak workload situations. In the moderate traffic scenarios, the lateral route clearances have higher frequencies during lower workload situations than during higher workload situations, again suggesting that the route clearances are used more often during these periods. However, altitude clearances seem to be issued periodically throughout the scenario, suggesting that they are not used specifically to minimize overall workload in the moderate traffic situations. One possible explanation to the data may be that regardless of the workload, the controllers used altitude clearances to proactively manage conflicts that are particularly complex. In general, complex conflicts in heavy congestion can require a significant time to craft a lateral solution. There also appear to be periodic oscillation patterns present in these data, but more analyses are needed to understand the exact nature of the oscillations.

Workload vs. Handoff-related Tasks. As stated earlier in the paper, the number of handoffs that a controller accepts from an upstream sector and initiates to a downstream sector is directly related to the number of aircraft in their sector. The average frequency of handoff initiation and acceptance per aircraft therefore shows considerable similarity between the high and moderate traffic/workload situations. The similarity between the two traffic levels suggest that the handoff-related tasks are performed similarly regardless of the traffic levels, suggesting that they are not omitted or delayed significantly during maximum workload situations unlike the bookkeeping tasks and the clearances described in the previous sections.

CONCLUSION

Based on the examination of how workload affects task distribution, the results suggest a qualitative shift in the types of tasks that controllers perform in low, moderate, and high workload states. During moderate traffic/workload states, controllers engaged in a significantly higher percentage of "bookkeeping" activities (e.g. datablock adjustments, etc.) than in the high workload states. If the workload reaches the maximum such that controllers need to manage their workload by selectively shedding or delaying tasks, it appears that they shed datablock adjustments and delay toggling/minimizing datablocks until they have enough time to attend to that task. None of these patterns emerge in the moderate traffic scenarios, presumably because controllers have enough mental resources to perform all of the tasks.

Since identifying task shedding strategies was not an explicit part of the training, it is likely that the task shedding behavior emerged as a natural extension of normal controller behavior. Controller participants have commented that in the high traffic scenarios, they were often in a "survival" and a "reactive" mode, in which they focused on keeping aircraft separated and abandoned activities related to providing service or maintaining a strategic plan for the traffic flow.

The bookkeeping tasks also served to maintain higher level of situation awareness to make sure the current status of the aircraft and that all necessary tasks were completed for the aircraft. When the overall workload is too high, maintaining this situation awareness cannot be offset by additional workload, resulting in less controller vigilance of the overall traffic situation.

Examining route vs. altitude clearances, we expected controllers to issue more altitude instead of lateral route clearances during peak workload situations because altitude clearances generally take less workload. Although the data appear to support this hypothesis, they were not conclusive.

Understanding how workload moderates task load distribution has significant potential for predicting true workload limits. If the pattern of delayed and dropped tasks show better consistency across controllers than the subjective workload ratings themselves, one can look for these patterns to indicate when the controllers are reaching their mental resource limits, which in turn could provide inputs to safety implications and capacity limits.

Acknowledgments. This study heavily leveraged previous simulation data from Distributed Air-Ground Traffic Management (DAG-TM) research, which was funded by Advanced Air Transportation Technologies (AATT) Project. The authors would like to thank members from the NASA Ames Airspace Operations Laboratory (AOL) in providing access to the data as well as continued support in this research. Finally, we would like give special thanks to Natalia Wehrle for her help in the data compilation/analyses.

REFERENCES

Athenes, S., Averty, P., Puechmorel, S., Delahaye, D., and Collet, C. (2002). ATC Complexity and Controller Workload: Trying to Bridge the Gap, HCI-Aero 2002: *International Conference on Human-Computer Interaction in Aeronautics*, Cambridge, MA.

Hilburn, B. G., Bakker, M. W. P., Pekela, W. D.,& Parasuraman, R. (1997). The effect of free flight on air traffic controller mental workload, monitoring and system performance. *In Proceedings of the 10th International CEAS Conference on Free Flight* (pp. 14–1/14–12).

Kopardekar, P. & Magyarits, S. (2003). Measurement and prediction of dynamic density, *5th USA/Europe Air Traffic Management R&D Seminar*, Budapest, Hungary.

Lee, P. U. (2005). A Non-Linear Relationship between Controller Workload and Traffic Count, *49th Annual Meeting of Human Factors and Ergonomics Society*, Orlando, FL.

Lee, P. U., Prevot, T., Mercer, J., Smith, N., & Palmer, E. (2005). Ground-side Perspective on Mixed Operations with Self-separating and Controller-managed Aircraft, *Proceedings of the 6th USA/Europe Air Traffic Management Research and Development Seminar*, Baltimore, MD.

Manning, C., Mills, S., Fox, C., Pfleiderer, E., and Mogilka, H. (2001). The Relationship Between Air Traffic Control Communication Events and Measures of Controller Taskload and Workload. *The Fourth International ATM R&D Seminar ATM-2001*, Santa Fe, NM.

Sperandio, J. C. (1978). The Regulation of Working Methods as a Function of Workload Among Air Traffic Controllers, *Ergonomics, Vol. 21, No. 3*, pp. 195-202.

Stein, E. (1985). Air Traffic Controller Workload: An Examination of Workload Probe, *DOT/FAA/CT-TN84/24. DOT/FAA Technical Center*. Atlantic City, NJ.