
10th ICCRTS – June 14th, 2005

Management of Multiple Human Supervisory Control Tasks



P.J. Mitchell and M.L. Cummings – MIT Humans and Automation Lab

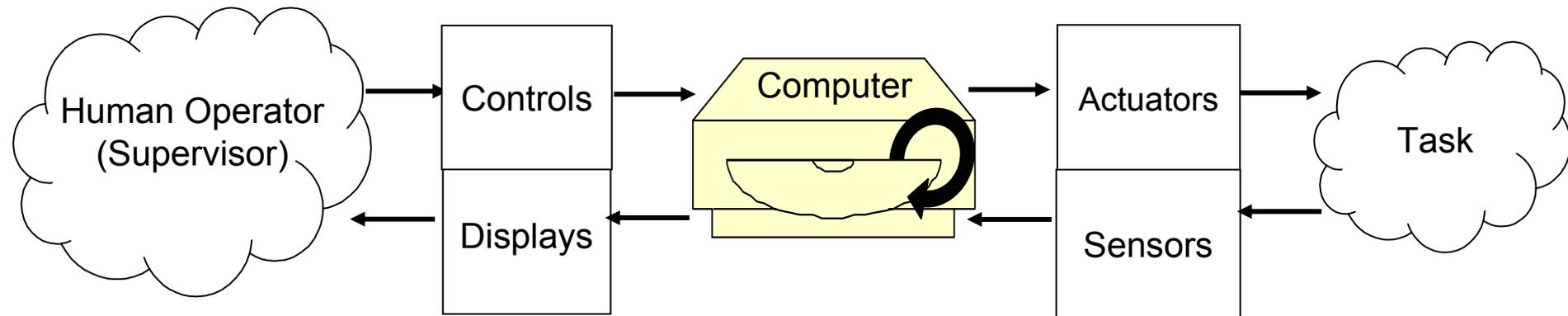


- **Introduction**
- **Motivation**
- **Background**
- **Simulation and Interface Design**
- **Hypotheses**
- **Methods**
- **Results and Discussion**
- **Conclusion**
- **Questions and Comments**



- **The five cyclical steps:**

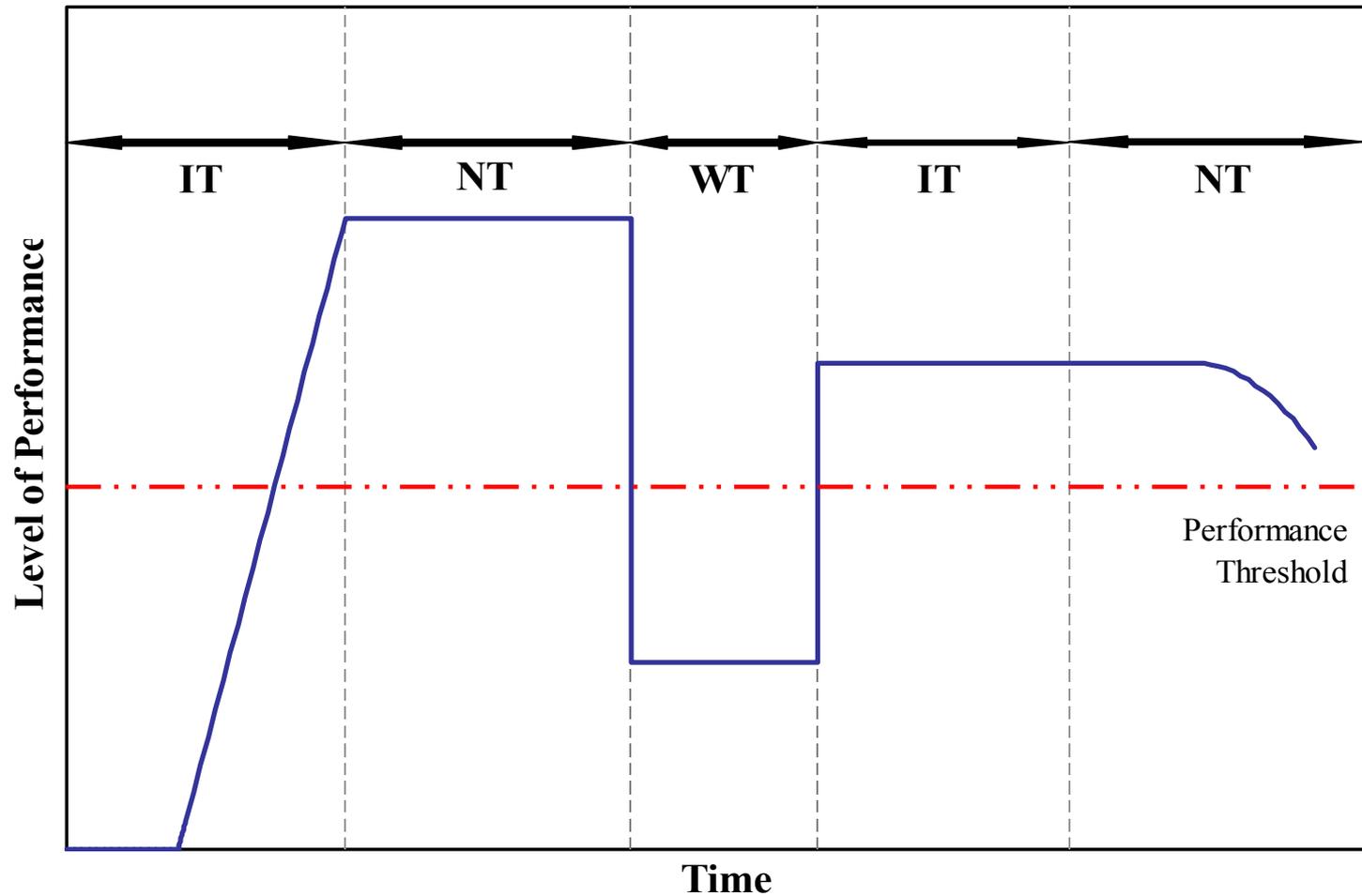
- **Planning** a computer-based task
- **Communicating** to the computer what was planned
- **Monitoring** the computer's actions for errors and/or failures
- **Intervening** when the plan has been completed or the computer requires assistance
- The human and computer **learn** from the experience



- **The role of multiple HSC tasks in network-centric operations**
 - HSC tasks are primarily cognitive in nature and generally do not require constant attention and/or control
- **Multiple HSC application of particular interest to military planners is operator supervision of a swarm of unmanned vehicles**
 - Want to maximize the number of unmanned vehicles a single operator can supervise, under time pressure
 - Major limiting factor is operator workload, which could drive system wait times
- **The role of automated decision support**
 - Not clear what type or level of decision support is appropriate for supervision of multiple vehicles

- **A framework for estimating the number of vehicles one human could control was proposed by Olsen and Wood (2004) with regard to traditional human-robot interactions**
- **Two main concepts in original work**
 - **Interaction Time (IT)**
 - **Neglect Time (NT)**
- **Fanout equation**
 - **# vehicles that a single operator can supervise = $NT/IT + 1$**
- **Missing an essential element**
 - **Wait Time (WT)**
 - **Dramatically impacts system performance and risk of failure in time-critical applications (e.g. C^2)**

Background – The Relationship Between IT, NT, and WT



- **Three main categories of wait times**

- 1. Interaction Wait Time (WTI)**

- **Interaction times that occur while the vehicle is in a degraded state**

- 2. Wait Time in the Queue (WTQ)**

- **Result from vehicles requiring attention simultaneously or near simultaneously**

- 3. Situation Awareness Wait Time (WTSA)**

- **Result from loss of situational awareness, operator does not realize vehicle is waiting**

Simulation and Interface Design – Navigation Display

The three major screen elements on the first display are:

- 1**
Map Display
- 2**
Mission Time
- 3**
Mission Planning & Execution

MauveControl v1.0 - Navigation

UAV 4
MISSION PLAN Legend

Weapons
Arm Payload Fire Payload

Movements
Move to Next Target Return to Base

Targets
Target Assignment

Queue	Name	BDA
1	T-4L	✓
2	T-11L	
3	T-5H	✓
4	T-1M	

Add Target Remove Target
T-11L

Way Points
Add WayPoint Remove WayPoint

Loiter Points
Add LoiterPoint Remove LoiterPoint

Mission Time
Current: 12:05:24 Z End: 12:30:00 Z
MISSION TIME Elapsed: 00:05:24 Remaining: 00:24:36

Map Display
A map showing a mission route with waypoints (T-1M, T-2H, T-3H, T-4L, T-5H, T-6H, T-7H, T-8H, T-9H, T-10M, T-11L, T-12L, T-13M, T-14H, T-15H, T-16H, T-17L, T-18M) and loiter points (1, 2, 3, 4). A green path highlights the current mission route. A large yellow circle is visible on the map.



Simulation and Interface Design – Decision Support Display

The four major screen elements on the second display are:

1

UAV Status

2

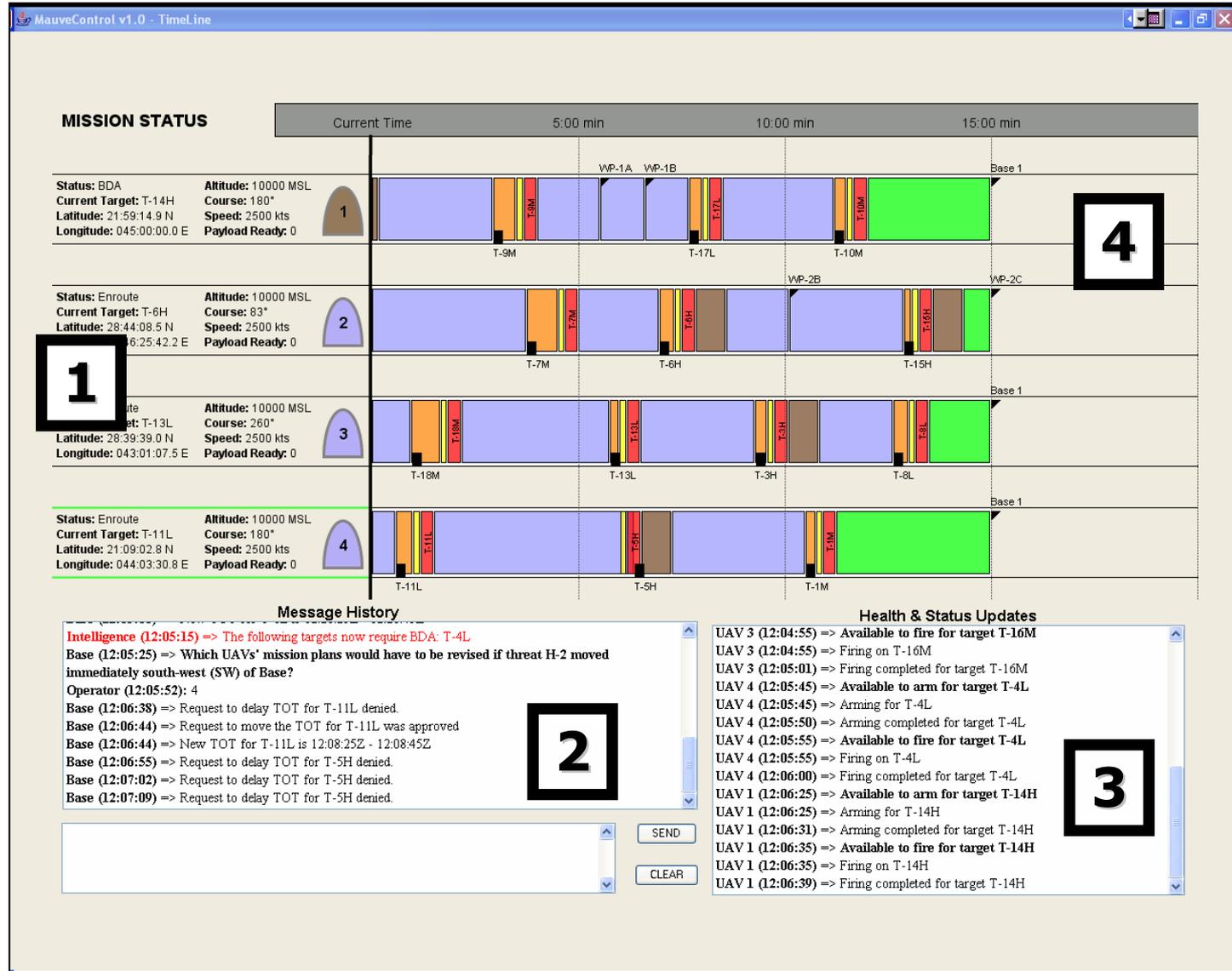
Chat Box

3

UAV Health & Status Updates

4

Decision Support



Methods – Experimental Apparatus



Primary: Level of Decision Support (Scheduling Assistance)

- **Between subjects**
- **Four levels**
 - 1. Manual**
 - 2. Passive**
 - 3. Active**
 - 4. Super Active**

Secondary: Amount of Schedule Re-planning (Operational Tempo)

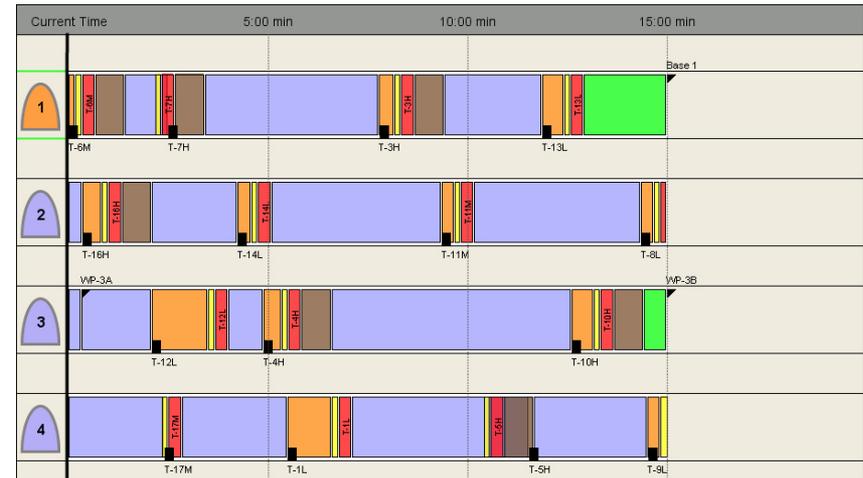
- **Within subjects**
- **Two levels, high and low**

Methods – Experimental Design – Types of Decision Support

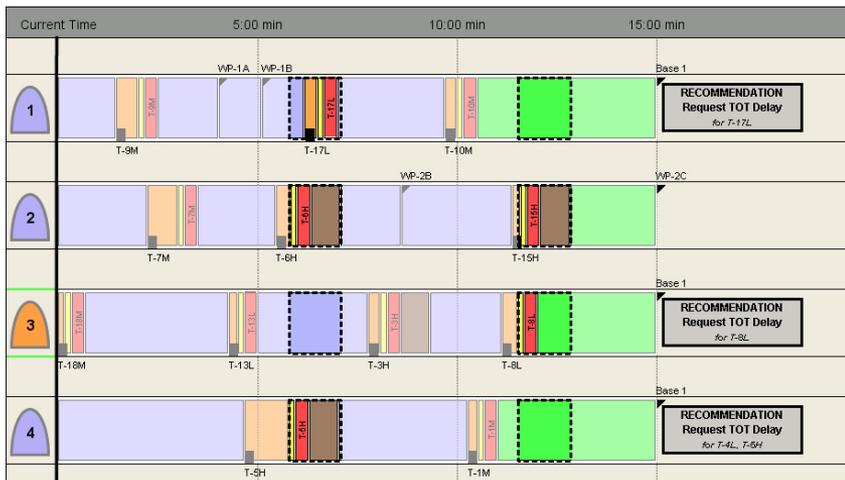
Manual

UAV	Current Target TOT / ETA	Next Waypoint or Loiterpoint	Next Expected Action	Upcoming Active Targets TOT / ETA			Mission Finish
1	T-6M 12:05:55Z - 12:06:15Z 12:05:34Z	-	Arming for T-6M 12:05:45 - 12:05:55	T-3H 12:13:55Z - 12:14:15Z 12:13:22Z	T-13L 12:18:10Z - 12:18:30Z 12:17:27Z	-	Base 1 12:23:36Z
2	T-16H 12:06:35Z - 12:06:55Z 12:05:55Z	-	Arming for T-16H 12:06:25 - 12:06:55	T-14L 12:10:20Z - 12:10:40Z 12:09:49Z	T-11M 12:15:25Z - 12:15:45Z 12:14:56Z	T-8L 12:20:25Z - 12:20:45Z 12:19:56Z	Base 1 12:23:00Z
3	T-12L 12:09:15Z - 12:09:35Z 12:07:41Z	WP-3A 12:05:48Z	Arming for T-12L 12:09:05 - 12:09:35	T-4H 12:11:05Z - 12:11:25Z 12:10:28Z	T-10H 12:16:55Z - 12:19:15Z 12:18:11Z	-	Base 1 12:26:19Z
4	T-17M 12:08:05Z - 12:08:25Z 12:07:53Z	WP-4A 12:06:46Z	Arming for T-17M 12:07:55 - 12:08:25	T-1L 12:10:55Z - 12:11:15Z 12:11:04Z	T-5H 12:16:10Z - 12:16:30Z 12:15:42Z	T-9L 12:20:35Z - 12:20:55Z 12:20:06Z	Base 1 12:25:19Z

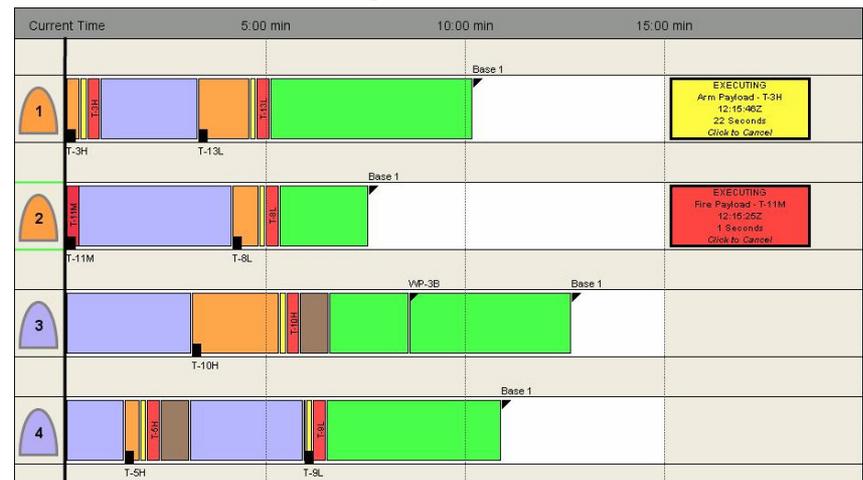
Passive



Active



Super Active



- **WTI**

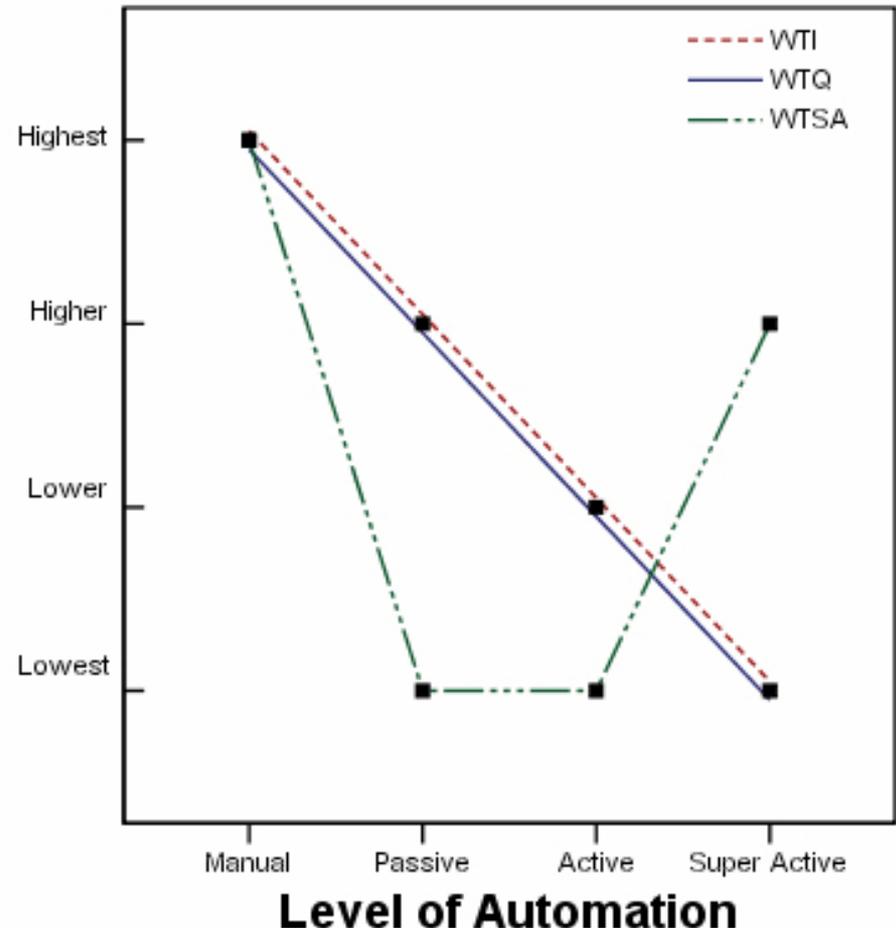
- **Should decrease with higher levels of automation**
 - **Visual timeline in all but manual level allows users to more easily determine inter-vehicle relationships**
 - **Recommendations in active level decrease planning time**
 - **Super active level eliminates execution time**

- **WTQ**

- **Follows same trend as WTI**

- **WTSA**

- **Situation awareness can decrease under both high and low workloads**

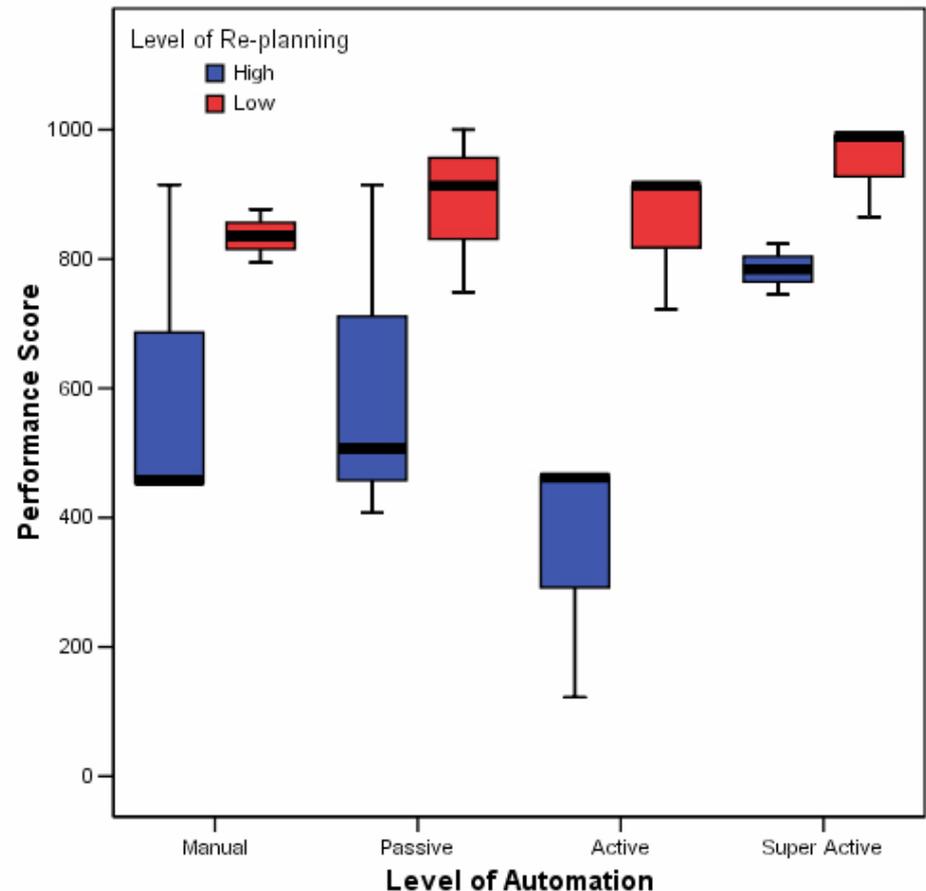


- **Human Performance Measures**
 - Performance Score
 - Number of Time on Target Delays Requested
- **System Performance Measures**
 - Wait Times
 - WTI, WTQ, WTSA
- **Situation Awareness**
 - Quasi-objective SA score on a 0-5 scale
 - Objective rating scales, but subjectively chosen by expert observer
- **Critical Events**
 - Number of times targets were *incorrectly* destroyed

**Level of re-planning significant
($p = 0.001$)**

**Level of automation marginally
significant ($p = 0.076$)**

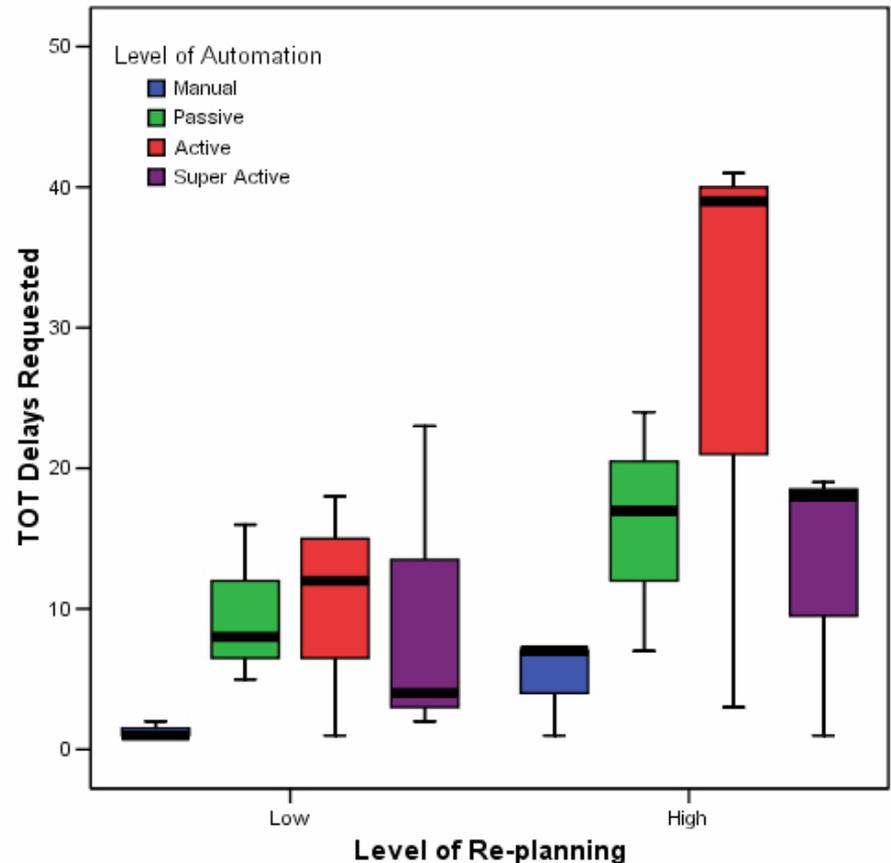
- **Low re-planning**
 - No difference between automation levels
- **High re-planning**
 - Super active had significantly higher performance scores than active ($p = 0.032$)
 - Poor performance of the active level is surprising in light of previous studies



Level of re-planning marginally significant ($p = 0.059$)

Level of automation marginally significant ($p = 0.096$)

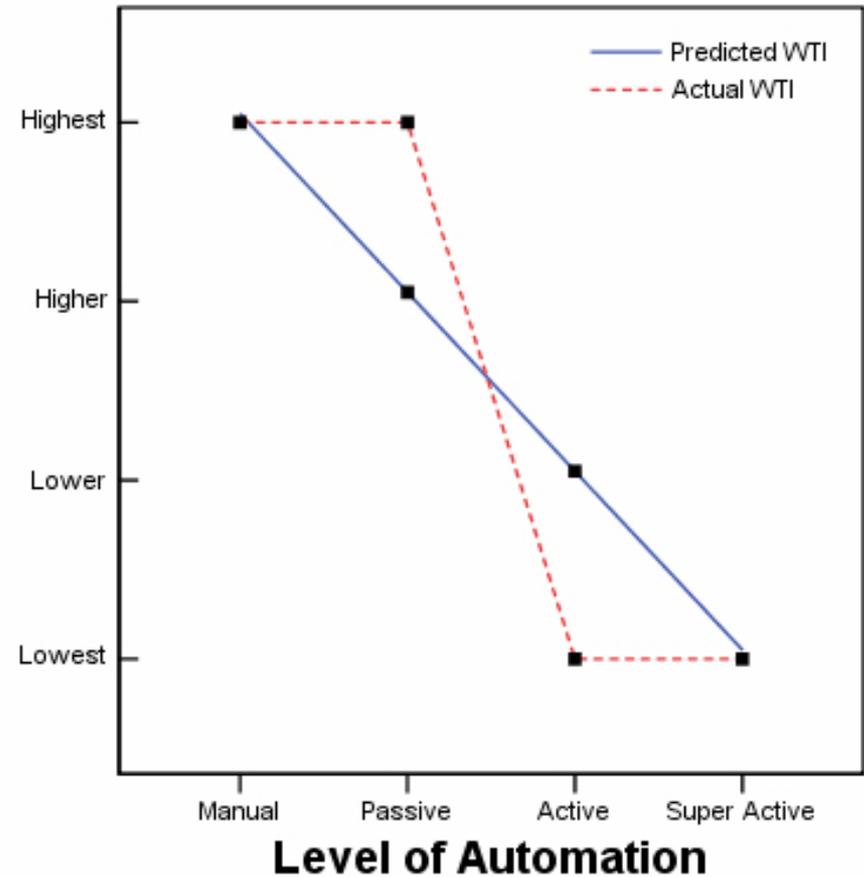
- **High re-planning**
 - **TOT requests under active level significantly higher ($p = 0.065$)**
 - **The driver for poor performance in the active level**
 - **Subjects in the active level tried to globally optimize their schedules and generally failed**



**Level of re-planning not significant
($p = 0.170$)**

**Level of automation significant
($p = 0.003$)**

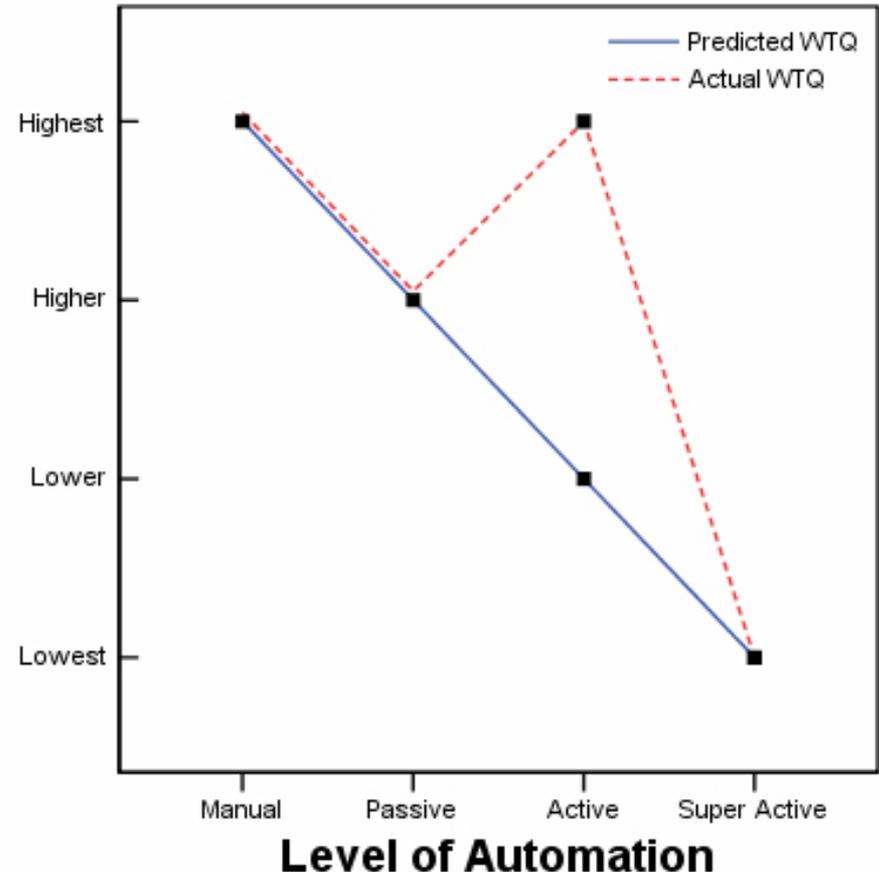
- **Significantly less total WTI for super active and active levels than manual and passive levels**
- **Reasonably consistent with expectations**



**Level of re-planning significant
($p = 0.001$)**

**Level of automation marginally
significant ($p = 0.063$)**

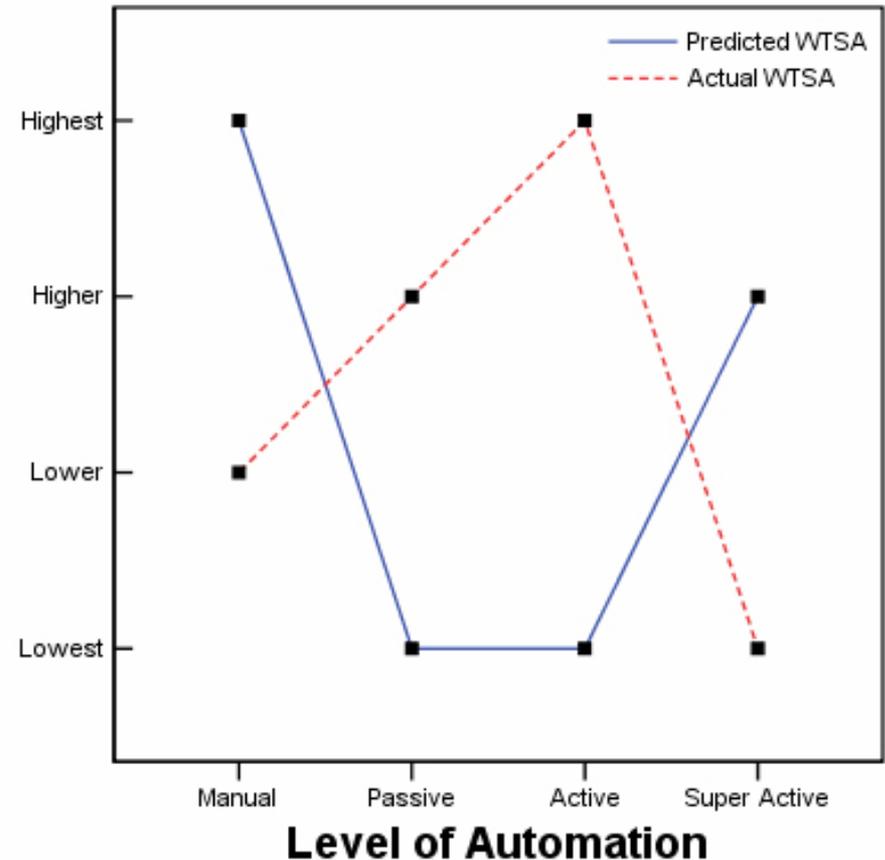
- **Consistent with expectations, except:**
 - **Under high re-planning, active level significantly higher WTQ than super active or passive, same as manual**
 - **Task queues built significantly higher in the active condition when under high workload**
 - **Due to extra time operators spent trying to adjust schedule**



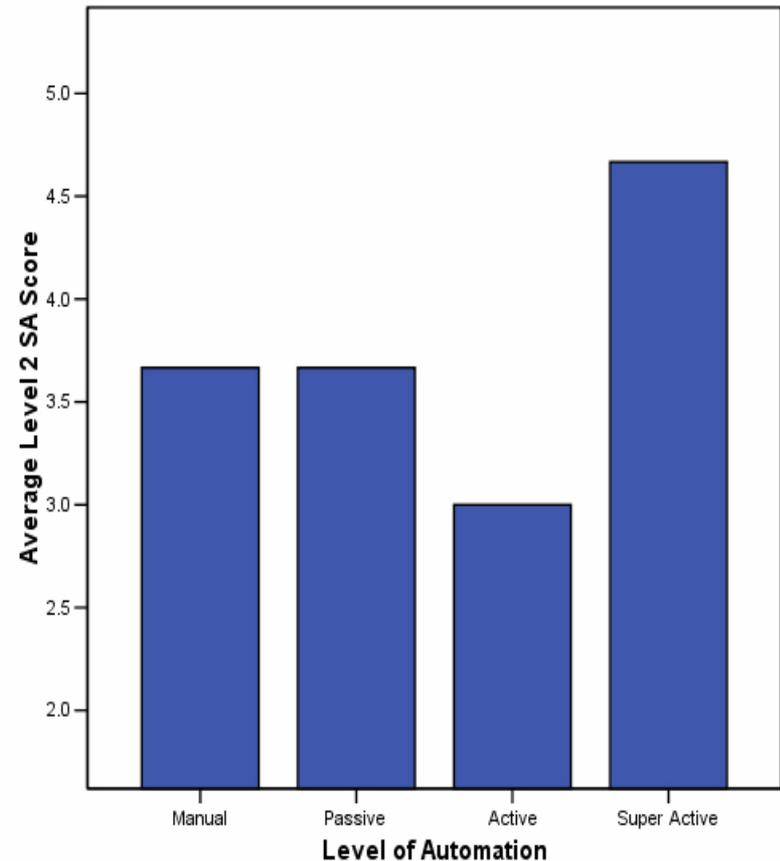
**Level of re-planning significant
($p = 0.001$)**

**Level of automation not significant
($p = 0.144$)**

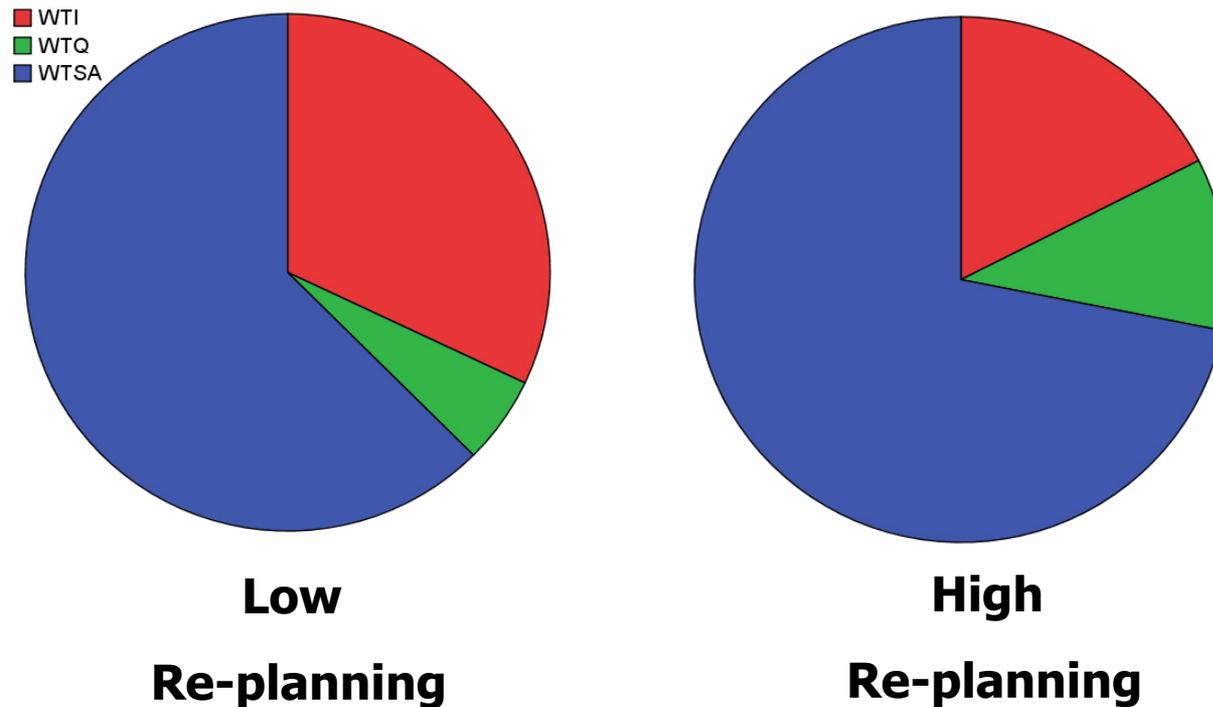
- **Super active significantly less WTSA than manual and active across both re-planning conditions**
- **Results not consistent with expectations**
 - **Inappropriate fixation on visual timeline**
 - **Further problems under active as users tried to globally optimize**



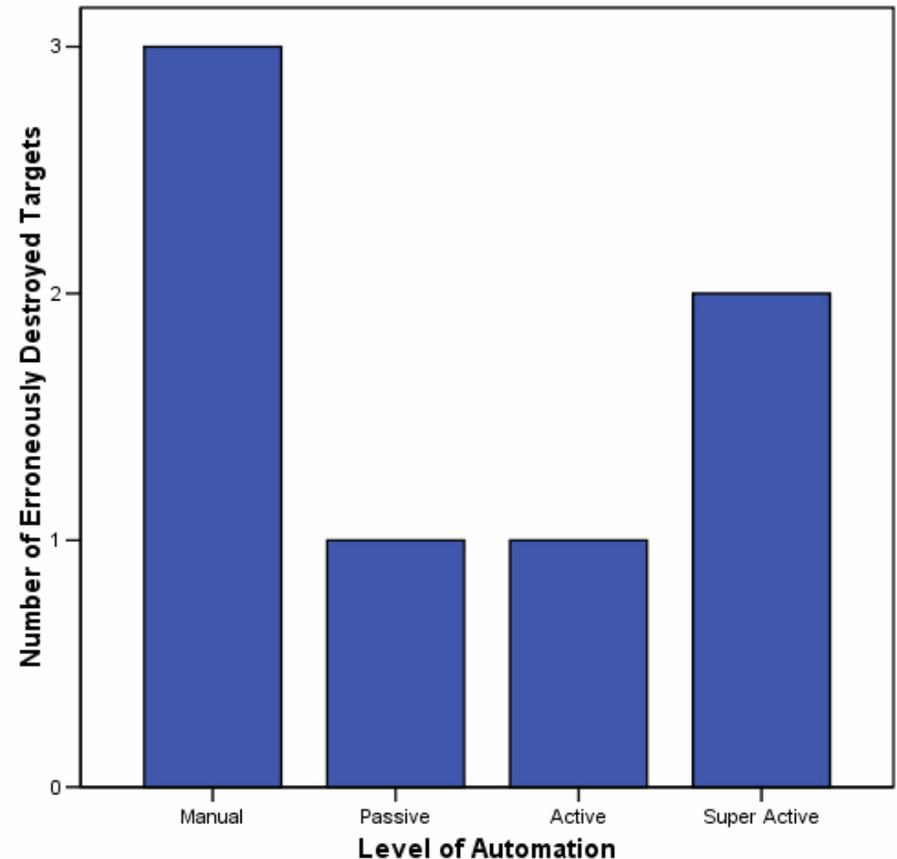
- **Overall**
 - **Level of automation not significant ($p=0.112$)**
 - **Same trends as WTSA**
- **Breakdown**
 - **SA is generally thought to have three levels (Endsley, 1995):**
 1. **Perception of important environmental cues**
 2. **Comprehension of the situation**
 3. **Projection of future events and dynamics**
- **Super active had significantly higher level 2 SA**
 - **More time to observe events on displays**
 - **No improvement in level three SA**



- **Total system wait time dominated by WTSA regardless of level of workload**
 - **WTI, WTQ can be reduced by greater autonomy**
 - **WTSA, which dominates total wait time, cannot be completely eliminated**



- **Occurred when operators erroneously fired upon targets**
 - **Friendly fire incidents, etc.**
- **Virtually no critical errors under low tempo condition**
- **Results suggest operators under manual and super active made more critical errors under high workload**



- **Super Active level of automation had the best human and system performance, but a higher number of catastrophic events**
- **Active level of automation unexpectedly produced the worst performance**
 - **Automation caused operators to attempt to globally optimize schedule, overwhelming them**
- **Passive level of automation best overall cost and performance benefits**
 - **A solid performer with no major drawbacks**
- **Total system wait time was dominated by wait time caused by lack of situation awareness**
- **Predictive model was good for WTI, WTQ, but not WTSA**
 - **Operators were never “under-loaded” due to number of vehicles being supervised**

MIT Humans and Automation Lab website

<http://halab.mit.edu>

Primary investigators' contact information

pmitchel@mit.edu

missyc@mit.edu

- Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1), 32-64.**
- Olsen, D. R., Jr., & Wood, S. B. (2004, April 24-29). *Fan-out: Measuring Human Control of Multiple Robots*. Paper presented at the Conference on Human Factors in Computing Systems, Vienna, Austria.**

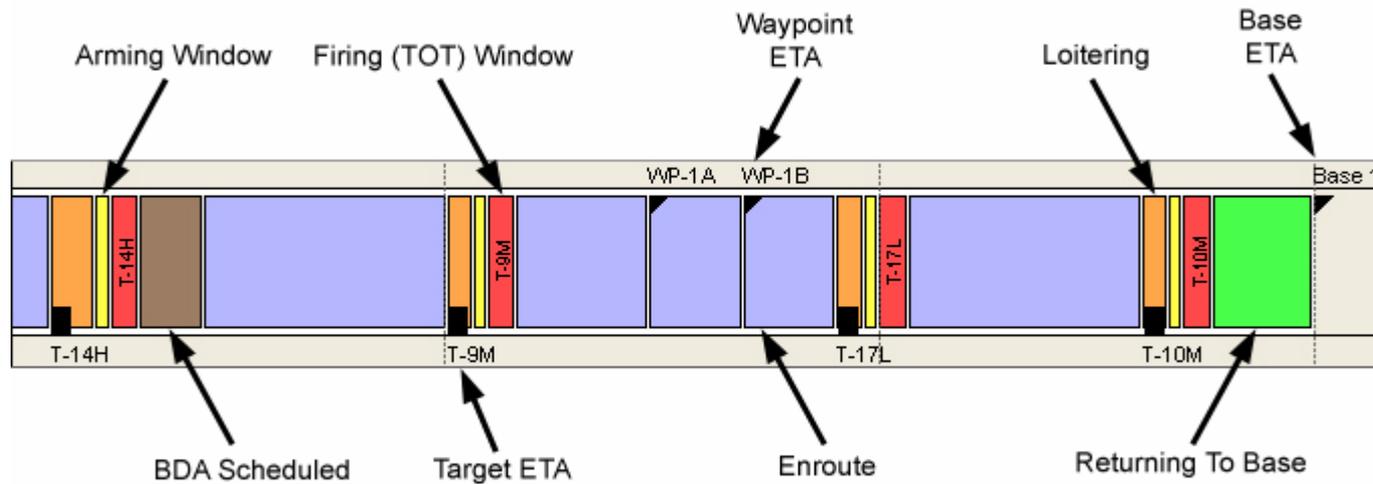
BACKUP SLIDES



Background – Levels of Automation

Automation Level	Automation Description
1	The computer offers no assistance: human must take all decision and actions.
2	The computer offers a complete set of decision/action alternatives, or
3	narrows the selection down to a few, or
4	suggests one alternative, and
5	executes that suggestion if the human approves, or
6	allows the human a restricted time to veto before automatic execution, or
7	executes automatically, then necessarily informs humans, and
8	informs the human only if asked, or
9	informs the human only if it, the computer, decides to.
10	The computer decides everything and acts autonomously, ignoring the human.

Simulation and Interface Design – Example Timeline

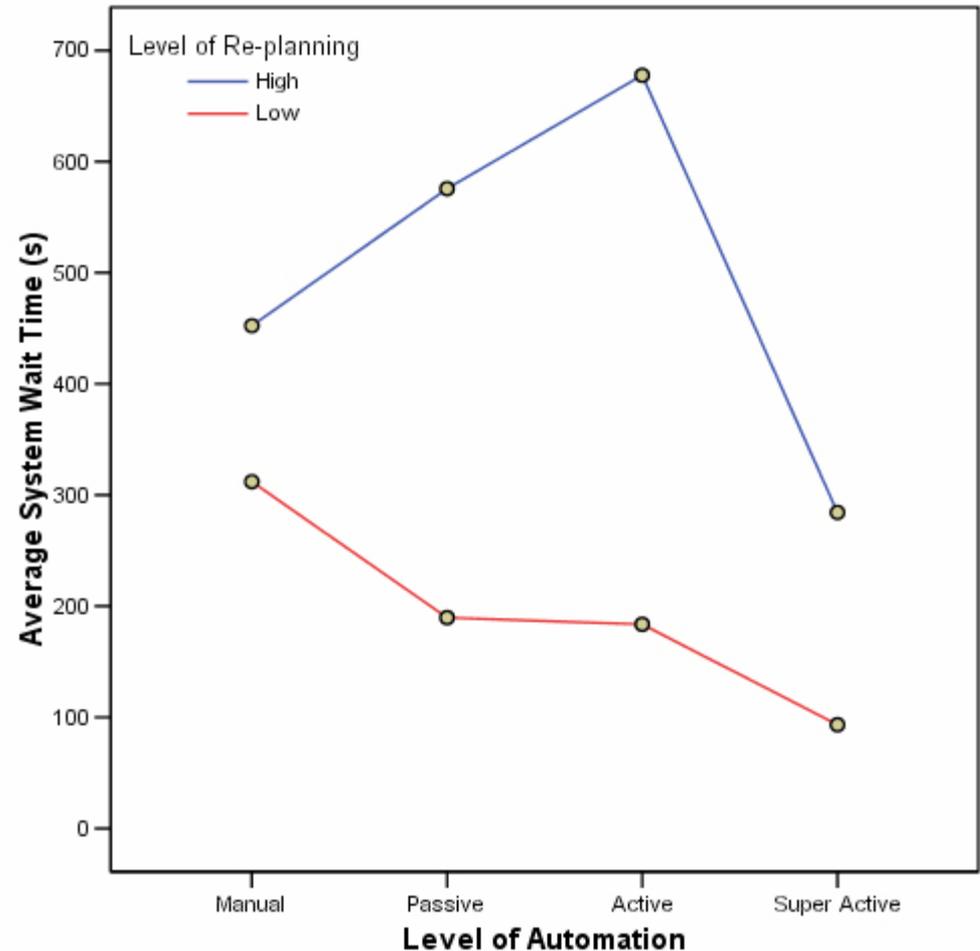


- **12 subjects total, 3 per automation level**
 - **Combination of MIT students, ROTC, and active duty military personnel**
 - **Average age = 26.3 years, 10 male, 2 female**
 - **9 are ROTC/Air Force officers**
 - **Mostly 2nd Lieutenants but up to Lieutenant Colonel**
 - **1-20 years experience, median 3**
 - **9 are pilots**
 - **Average flight hours 120**
 - **2 had previous (small) UAV experience**

**Level of re-planning significant
($p < 0.001$)**

**Level of automation significant
($p = 0.018$)**

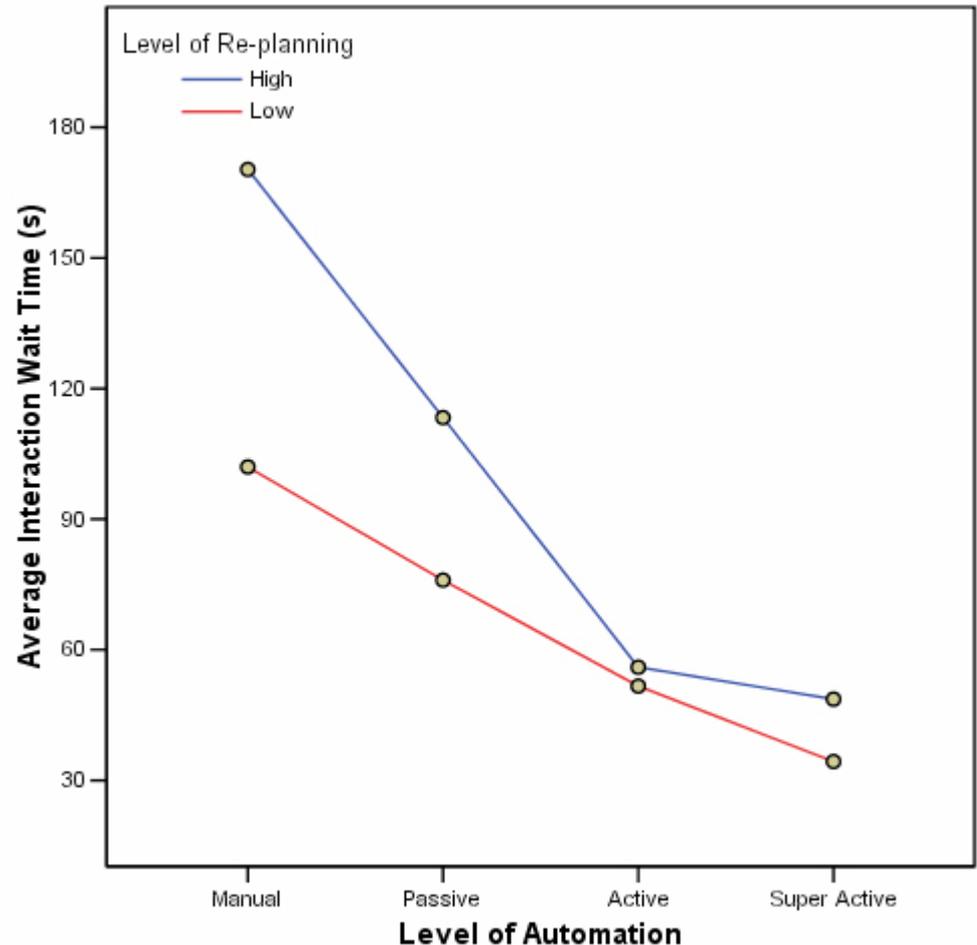
- **Super active WTT significantly less than all other automation levels**
- **High re-planning only**
 - **Active level significantly higher than super active**



Level of re-planning not significant ($p = 0.170$)

Level of automation significant ($p = 0.003$)

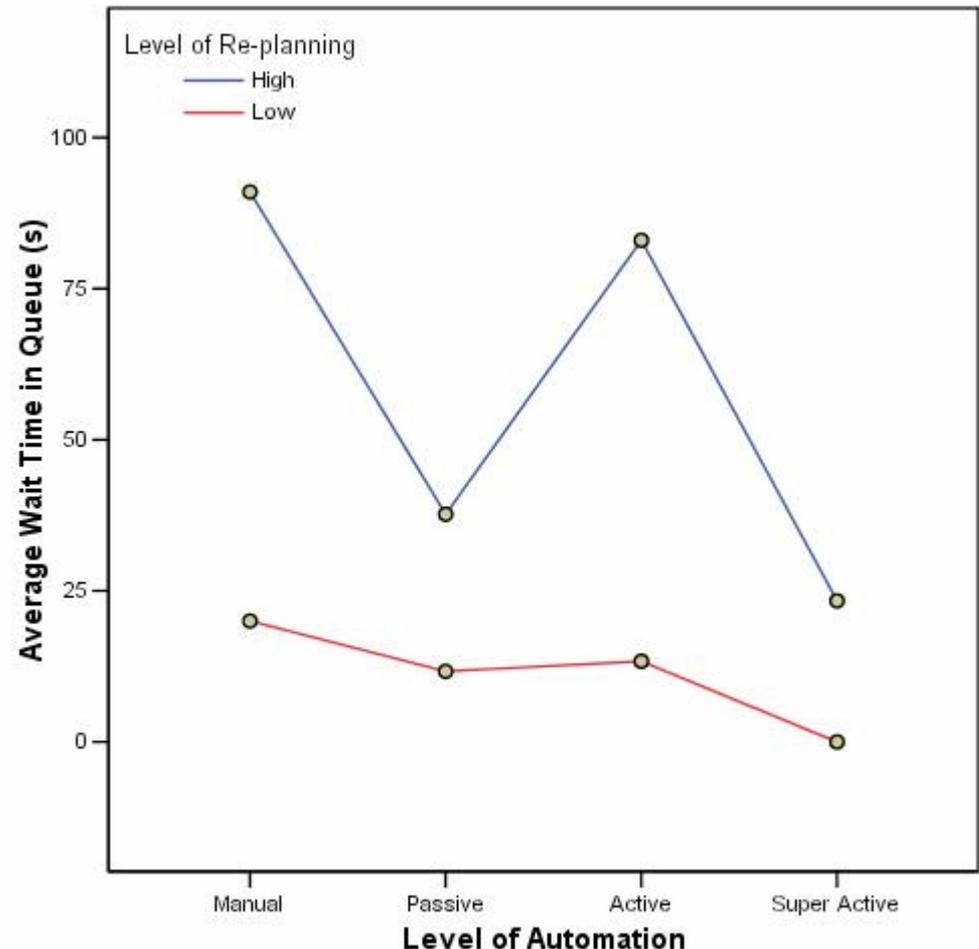
- **Significantly less for super active and active levels than manual and passive levels**



**Level of re-planning significant
($p = 0.001$)**

**Level of automation marginally
significant ($p = 0.063$)**

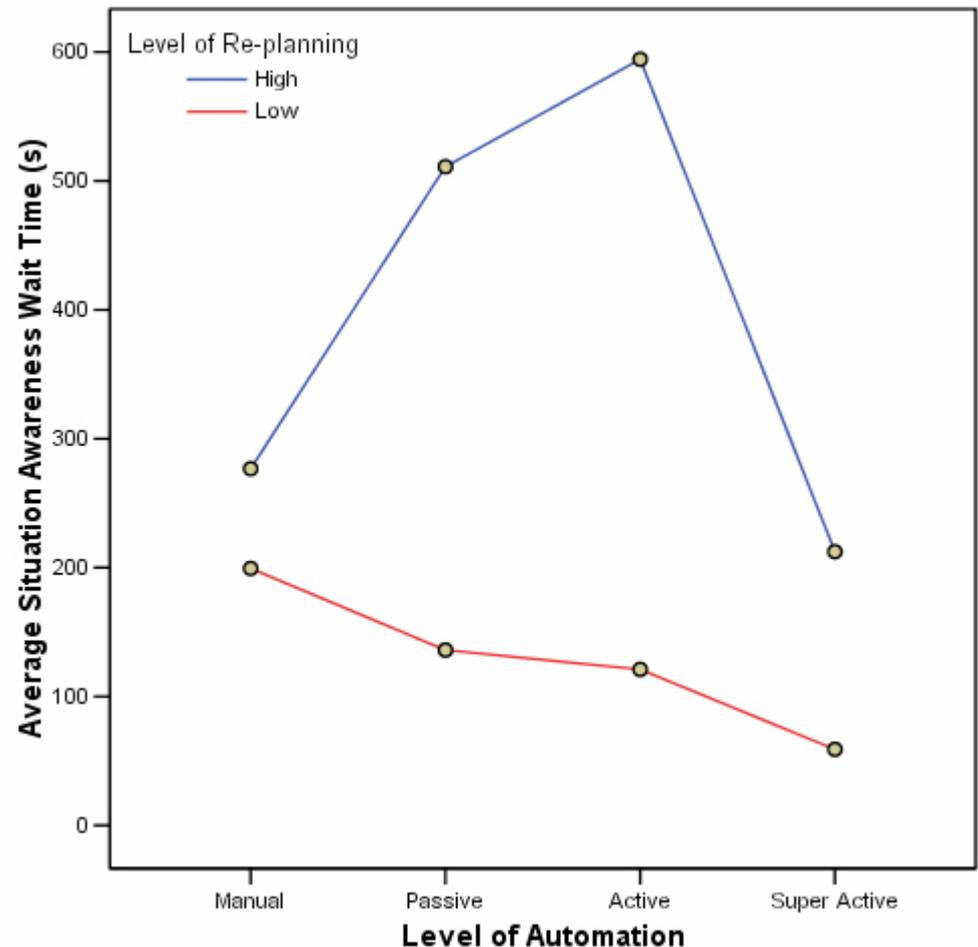
- **High re-planning only**
 - **Active level significantly higher WTQ than super active or passive, same as manual**



**Level of re-planning significant
($p = 0.001$)**

**Level of automation not
significant ($p = 0.144$)**

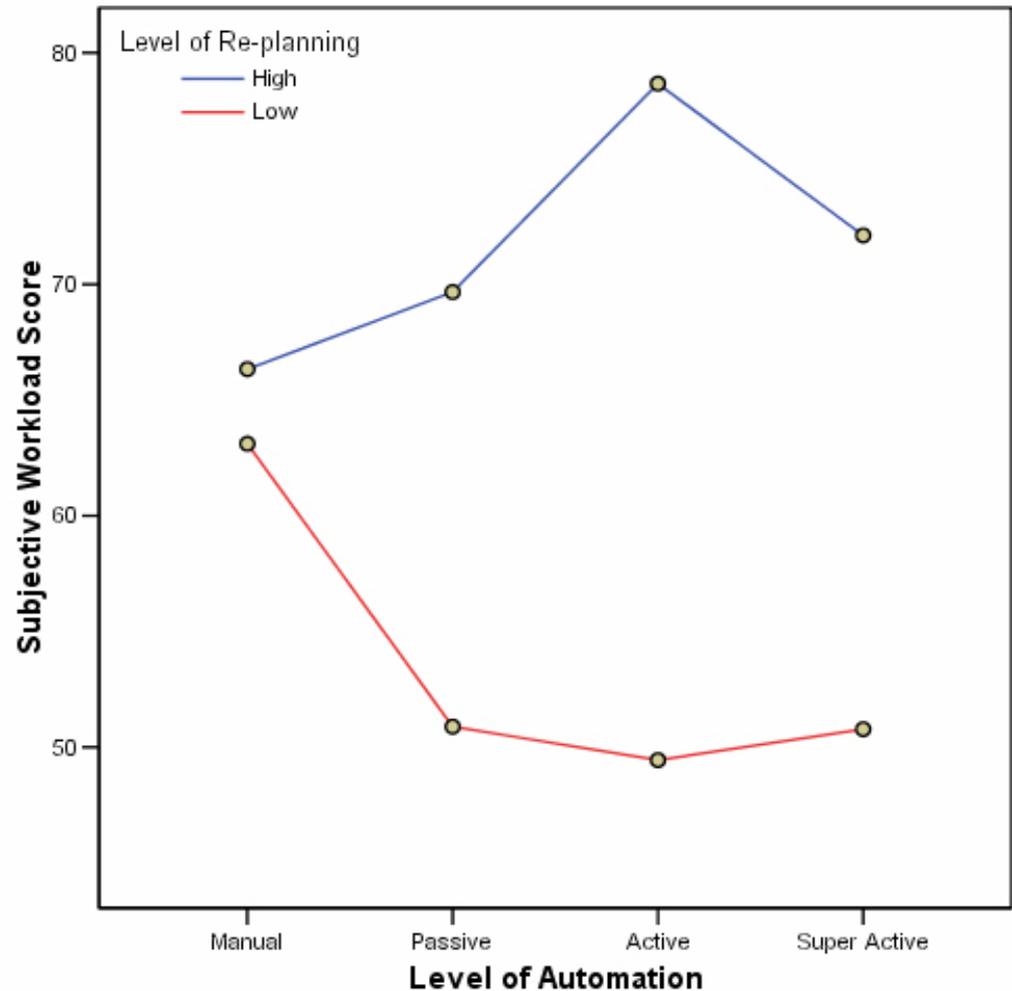
- **Cell comparisons**
 - **Super active significantly less than manual and active across both conditions**
 - **HRP, active significantly higher than manual and super active**



Level of re-planning significant ($p < 0.001$)

- No significant difference in subjective workload for the manual level under different operational tempos

Level of automation not significant ($p = 0.779$)

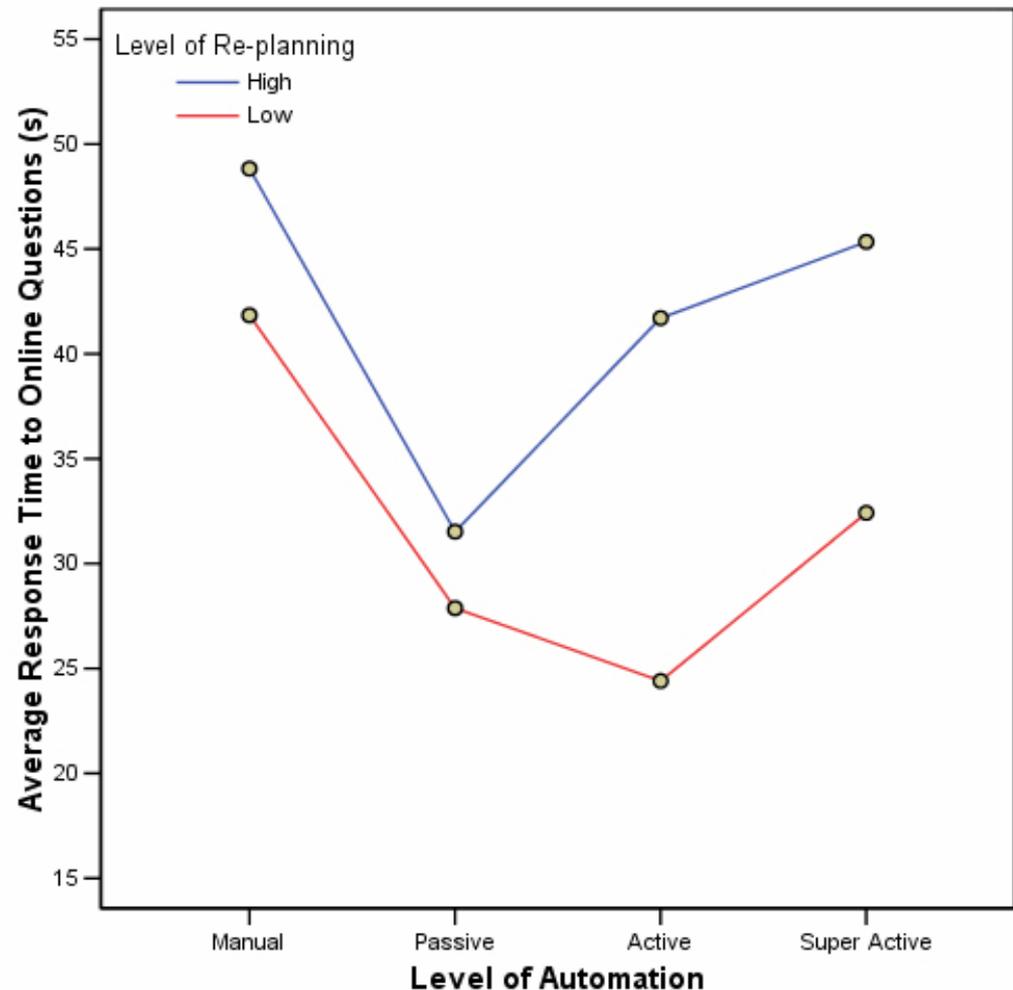


Level of re-planning significant (p=0.003)

- Spare capacity of the manual and passive levels did not change across re-planning conditions

Level of automation significant (p=0.002)

- Manual and super-active automation levels statistically have the same spare capacity, which is lower than active and passive levels
 - Manual = workload
 - Super Active = low SA



Results and Discussion – Best and Worst Performers

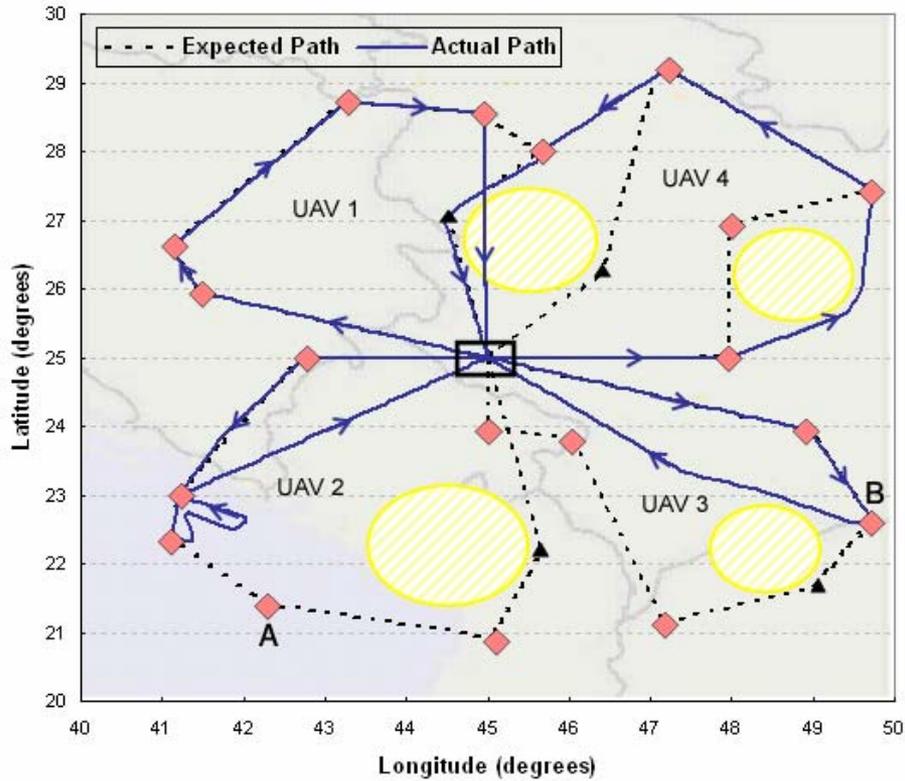
Dependent Variable	Level of Automation	
	Worst Performer	Best Performer
Performance Score	Active	Super Active
Number of TOT Delay Requests	Active	Manual
Interaction Wait Time (WTI)	Manual / Passive	Super Active / Active
Wait Time in the Queue (WTQ)	Active / Manual	Super Active / Passive
Situation Awareness Wait Time (WTSA)	Active	Super Active
Situation Awareness (SA)	Active	Super Active
Erroneous Target Critical Events	Super Active / Manual	Active / Passive
Subjective Workload	Active / Manual	Super Active / Passive
Secondary Workload	Manual	Passive

Results and Discussion – Summary of Main Effects (p-values)

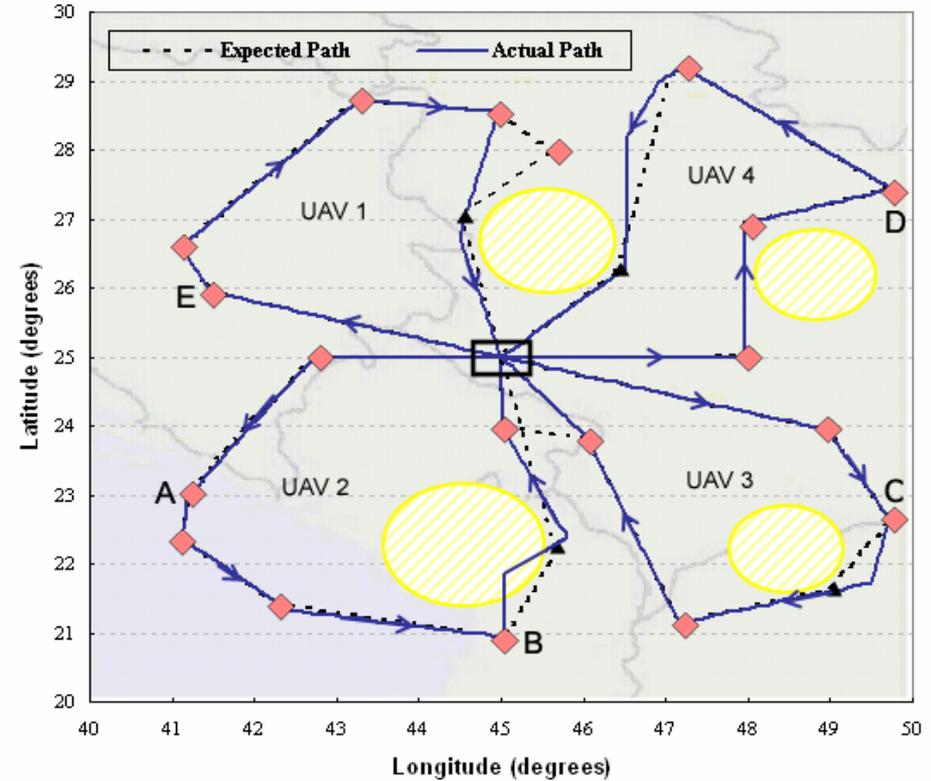
Dependent Variable	Independent Variable	
	Level of Automation	Level of Re-planning
Performance Score	0.076	0.001
Number of TOT Delay Requests	0.096	0.059
Interaction Wait Time (WTI)	0.003	0.170
Wait Time in the Queue (WTQ)	0.063	0.001
Situation Awareness Wait Time (WTSA)	0.144	0.001
Situation Awareness (SA)	0.112	0.002
Erroneous Target Critical Events	0.878	0.034
Subjective Workload	0.779	<0.001
Secondary Workload	0.002	0.003

LEGEND ($\alpha=0.05$)	
	Significant Result
	Marginally Significant Result

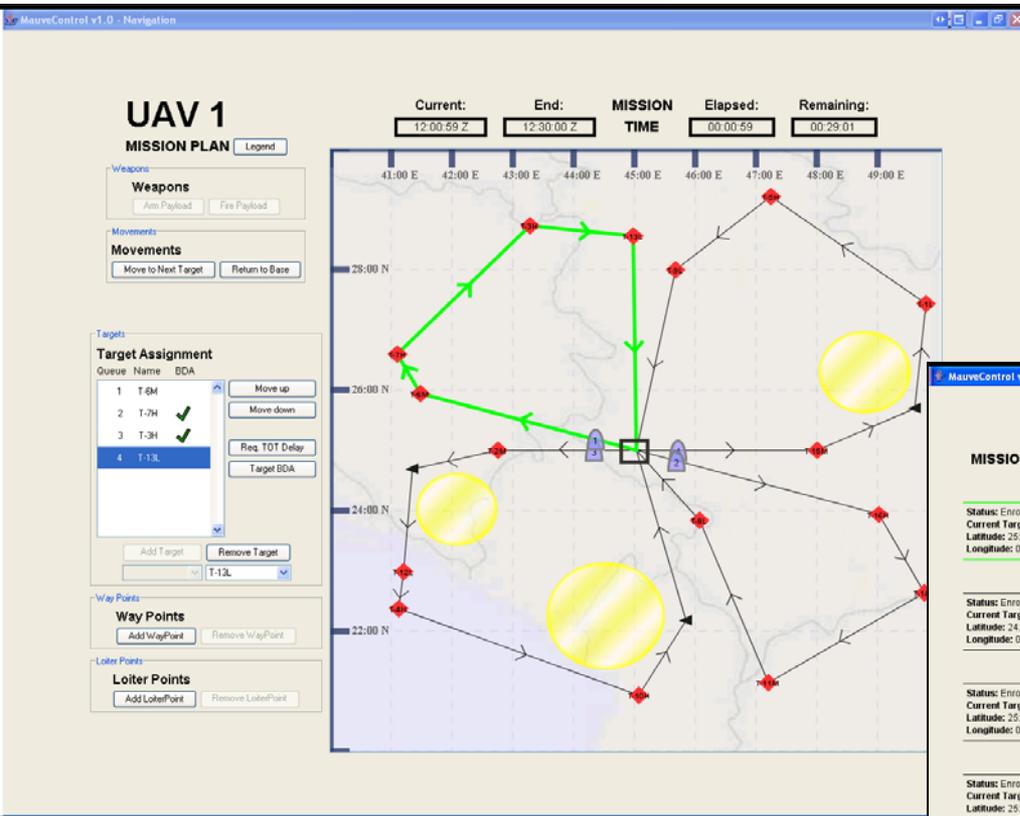
Results and Discussion – Overloaded Operator Coping Strategies



Cognitive Shedding

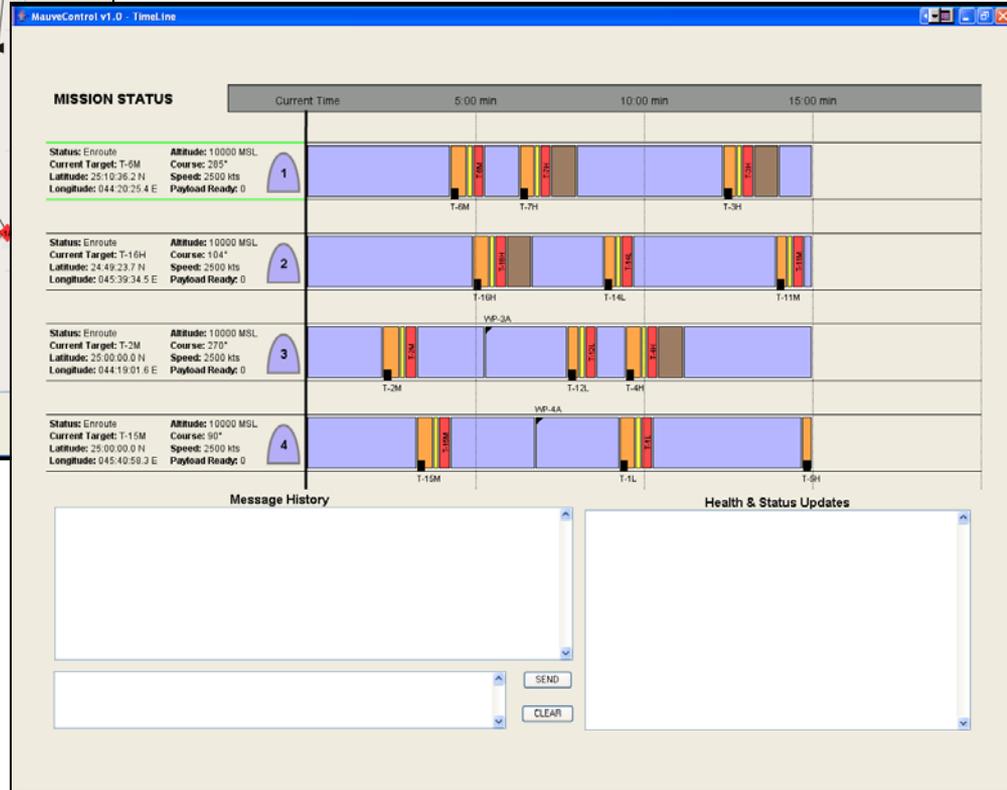


Degraded Level of Management



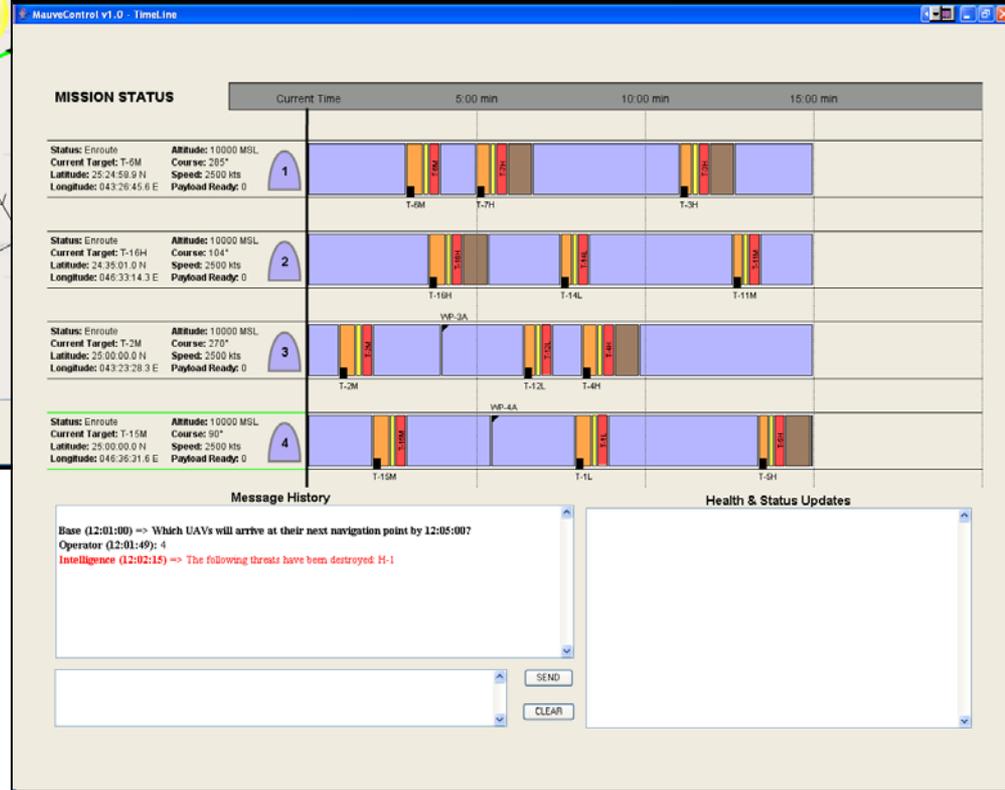
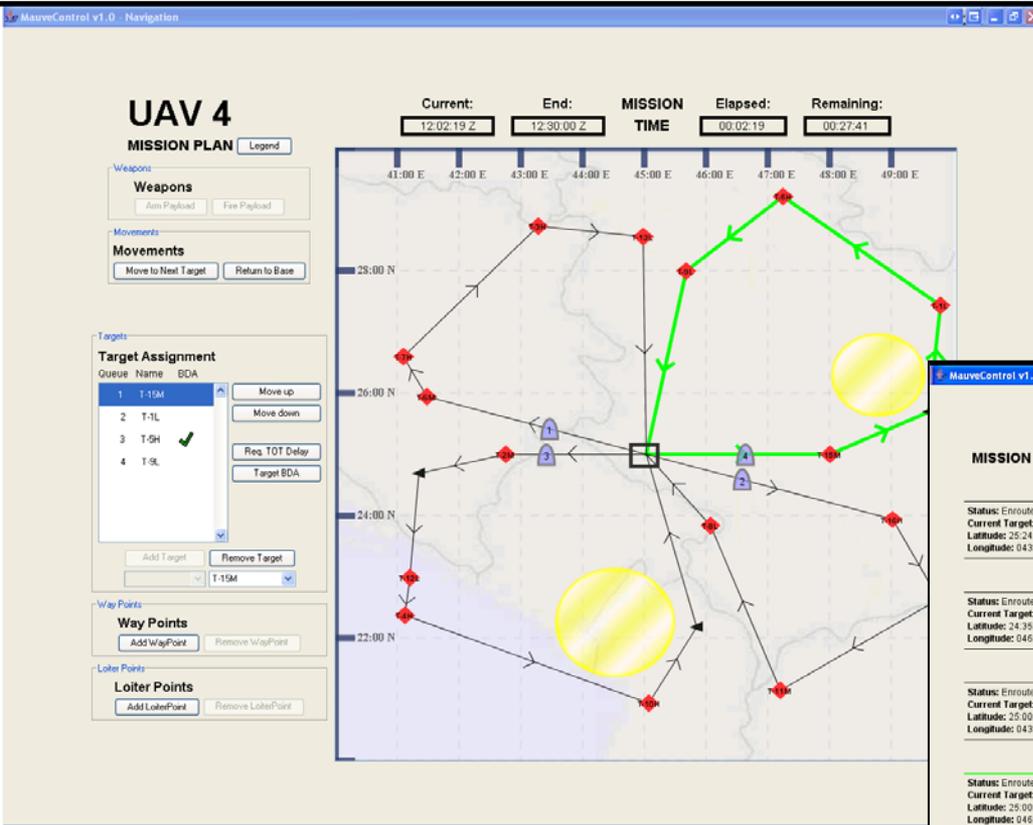
Super-Active Automation High Re-planning Scenario

12:00:59



Results and Discussion – Example Test Session – 2

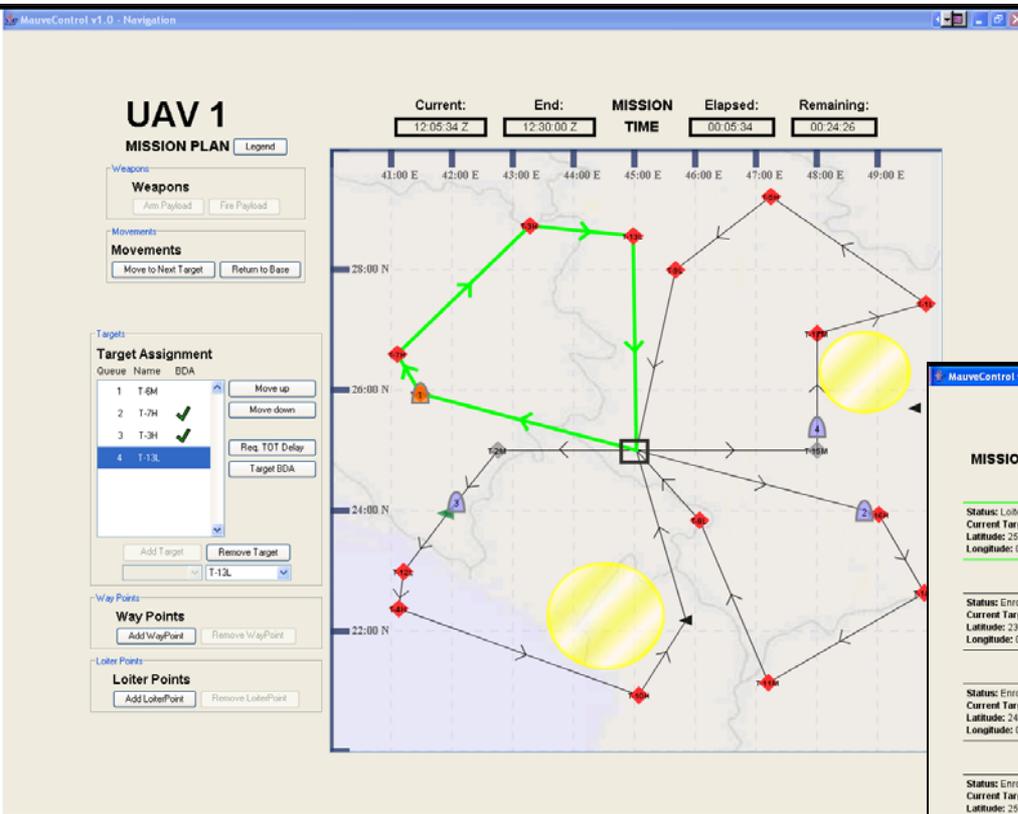
Super-Active Automation High Re-planning Scenario



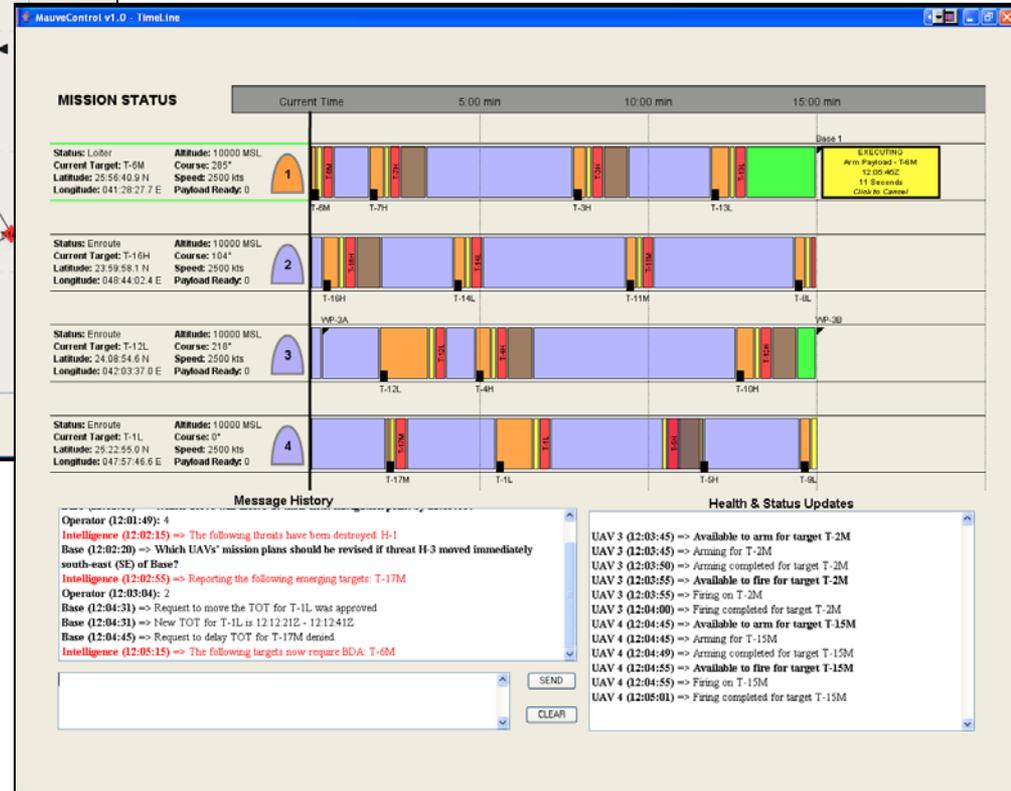
12:02:19



Super-Active Automation High Re-planning Scenario

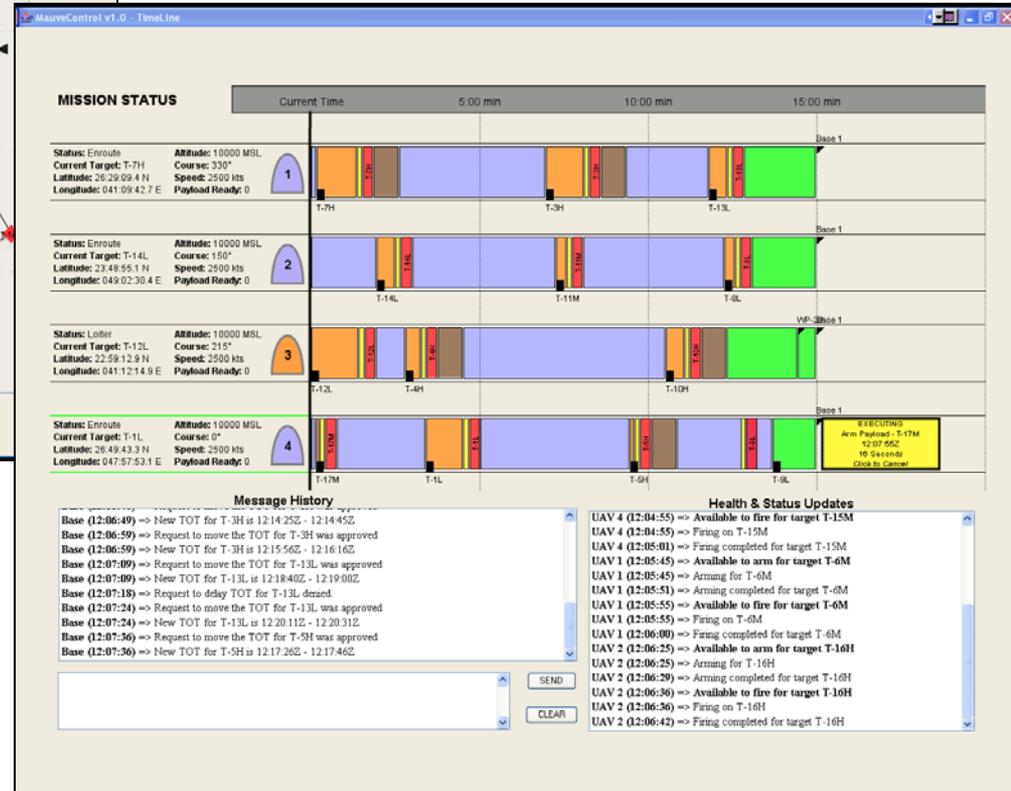
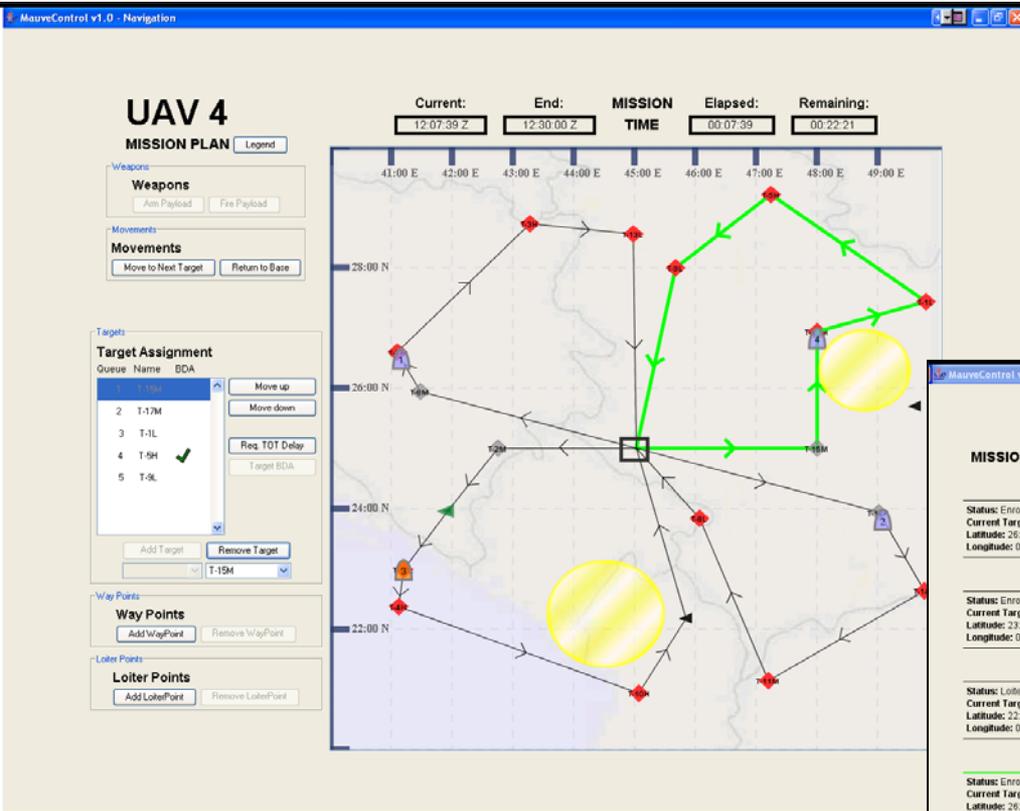


12:05:34



Results and Discussion – Example Test Session – 4

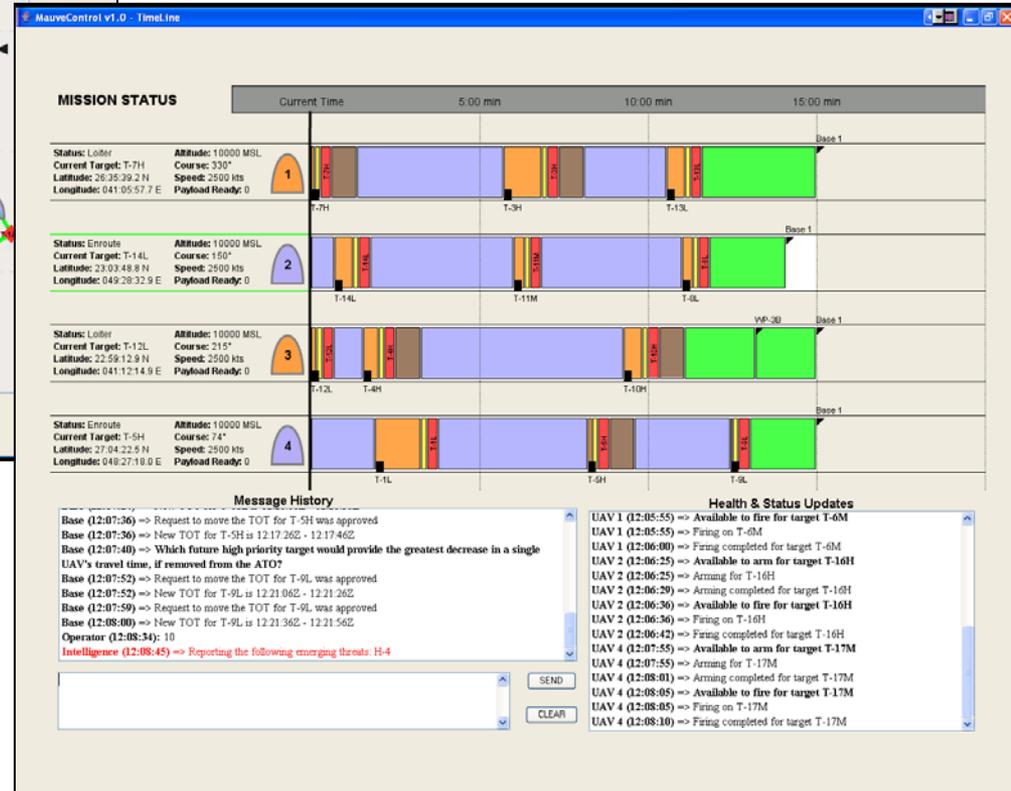
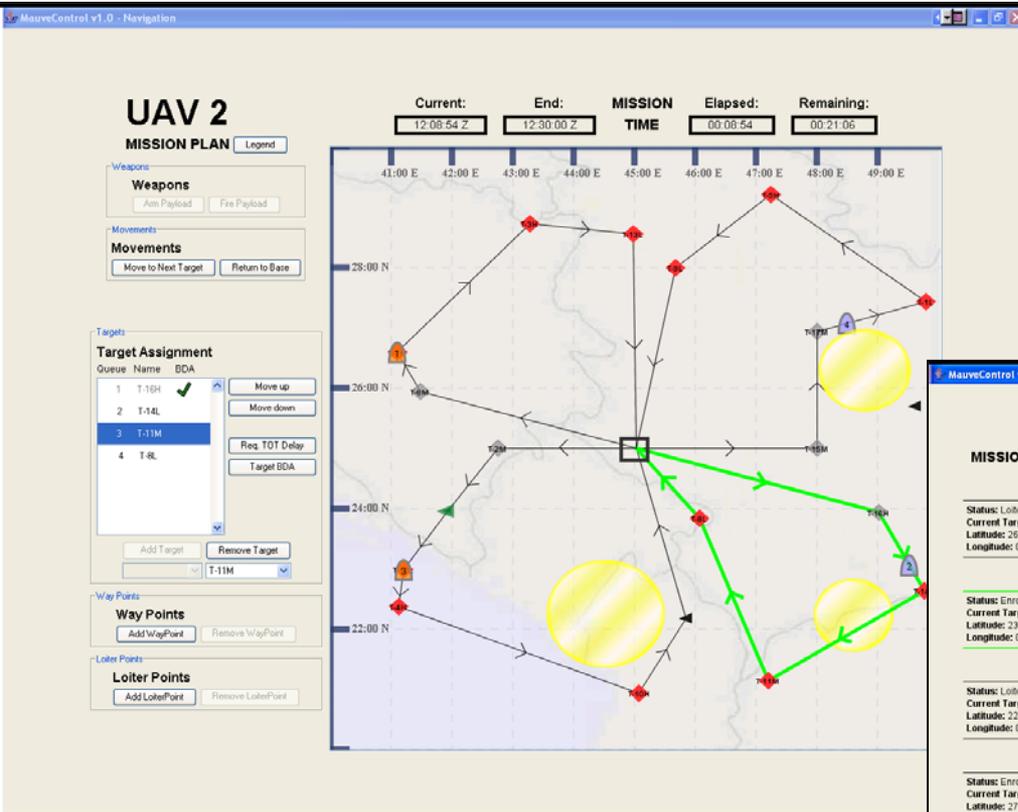
Super-Active Automation High Re-planning Scenario



12:07:39



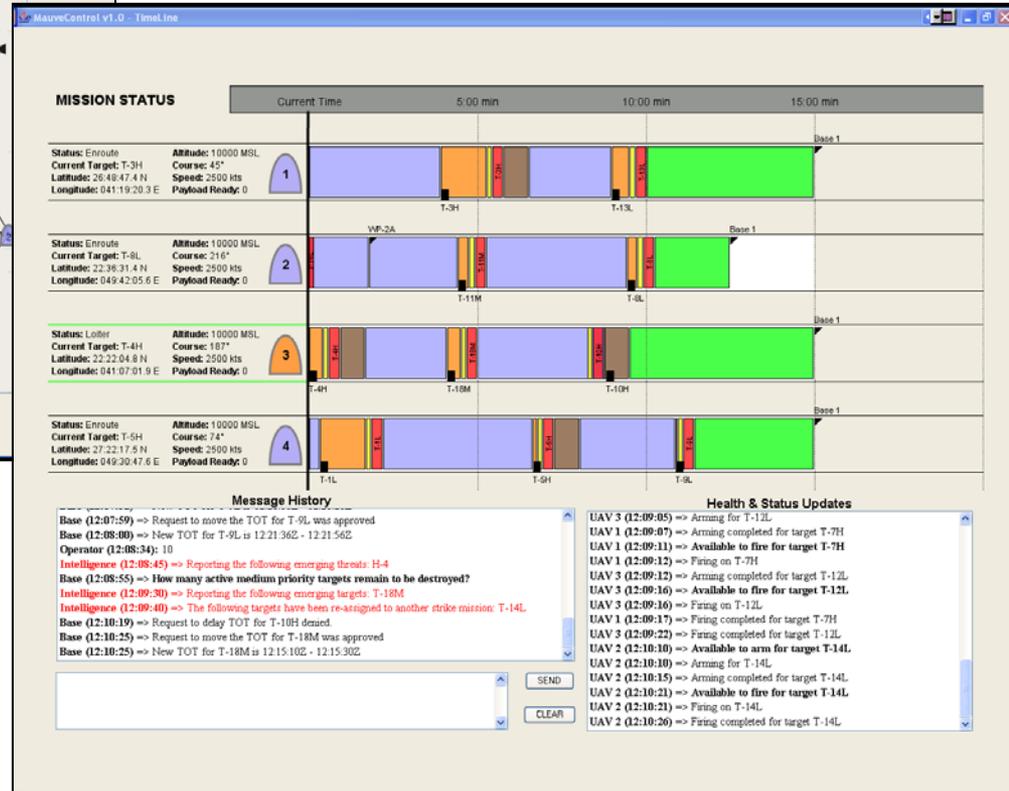
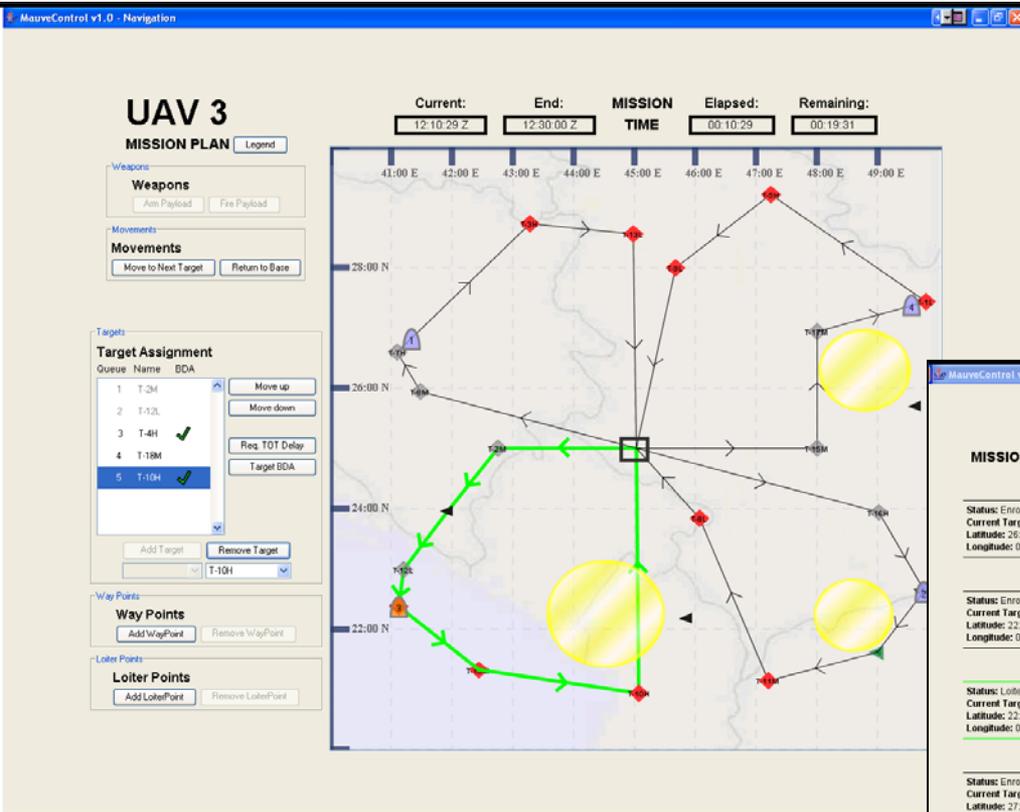
Super-Active Automation High Re-planning Scenario



12:08:54



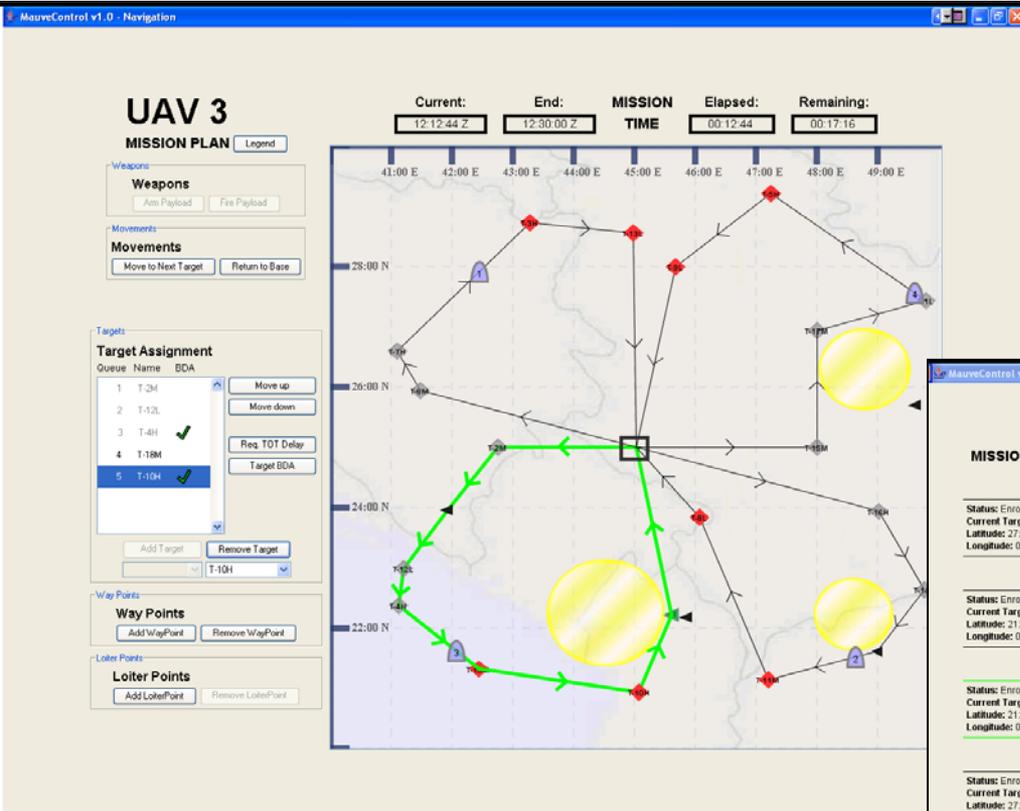
Super-Active Automation High Re-planning Scenario



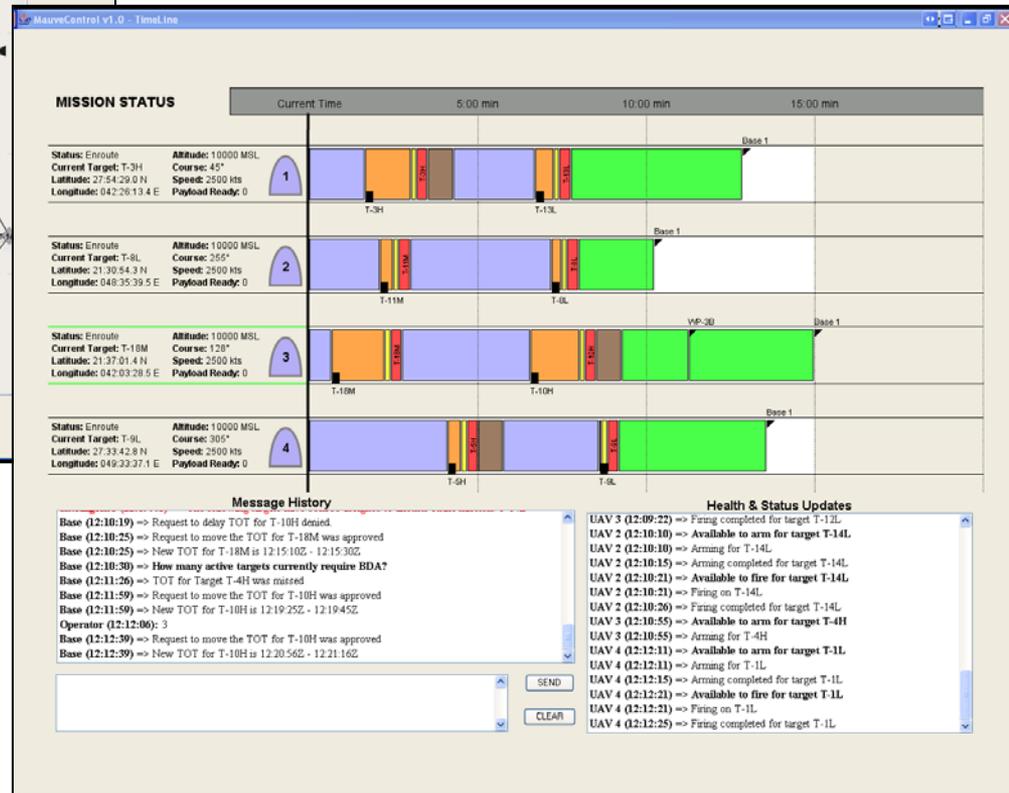
12:10:29



Results and Discussion – Example Test Session – 7

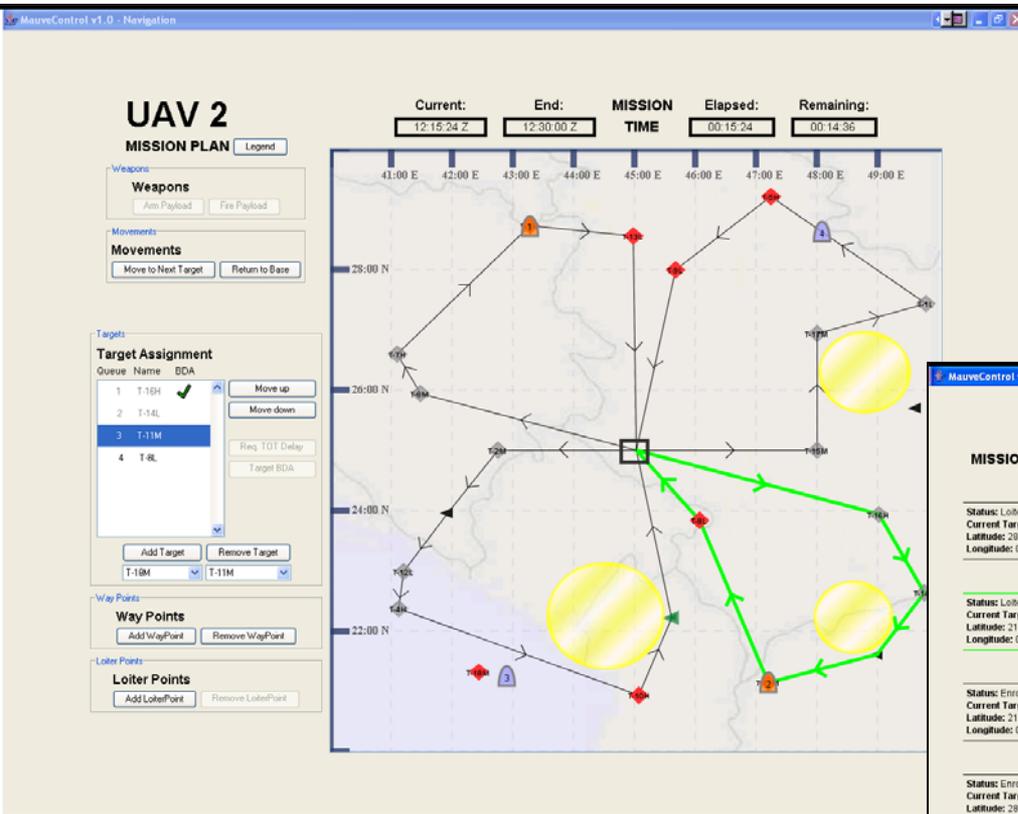


Super-Active Automation High Re-planning Scenario

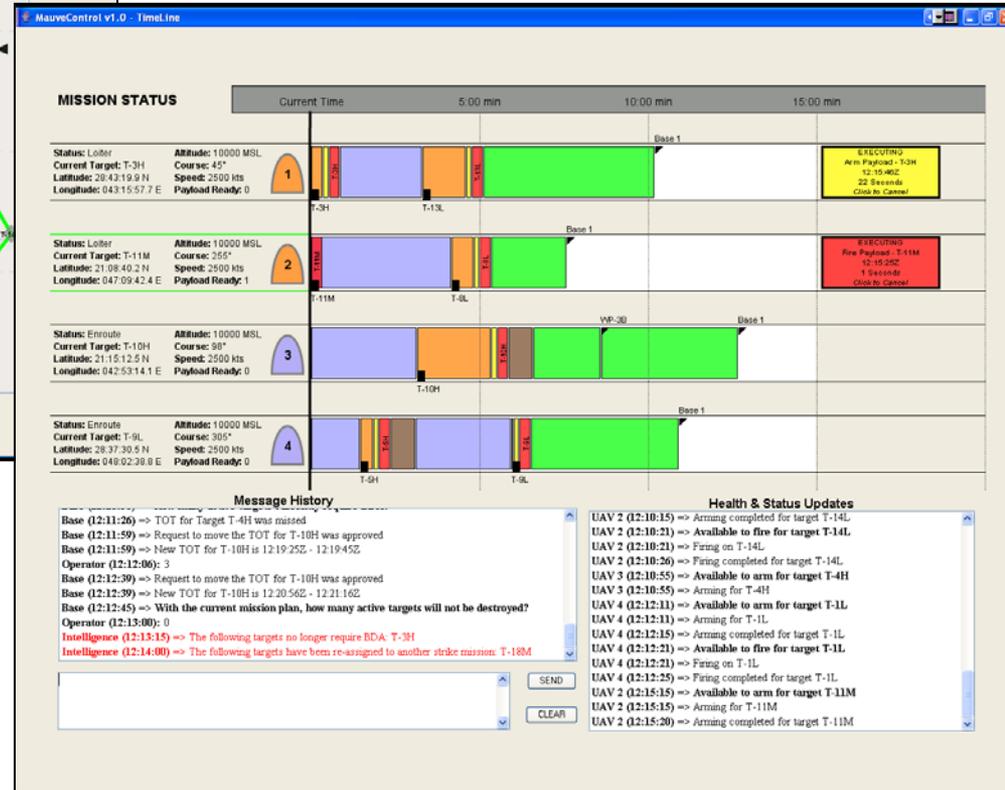


12:12:44





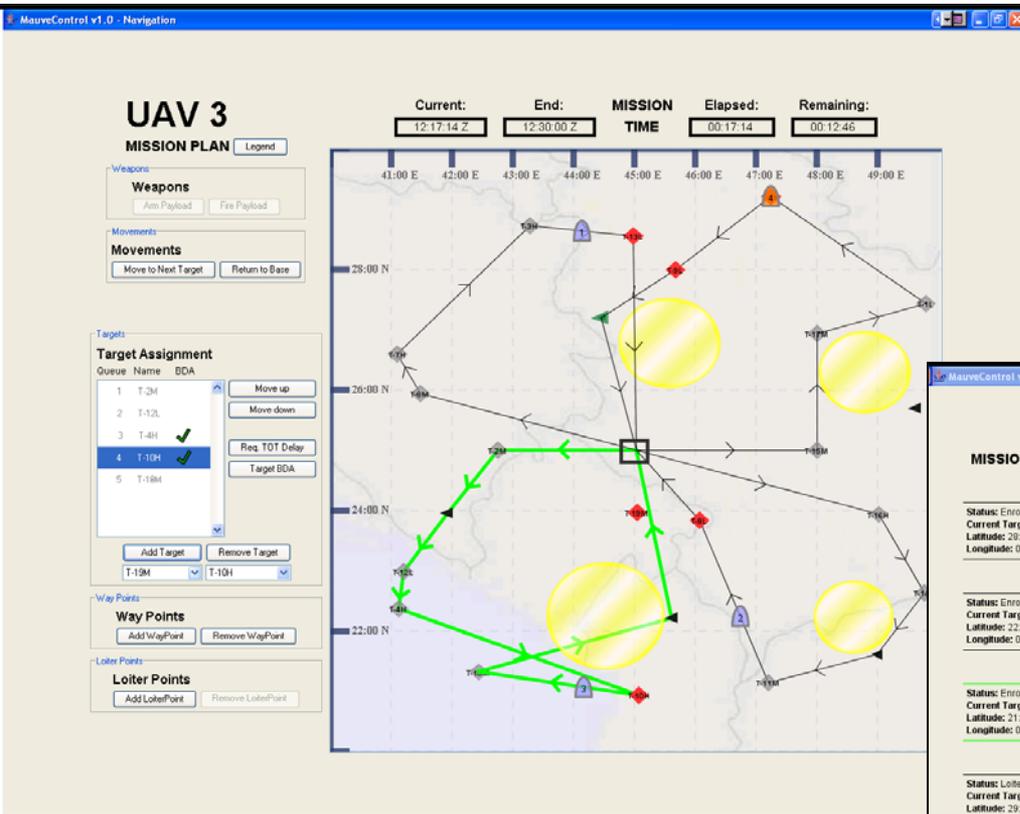
Super-Active Automation High Re-planning Scenario



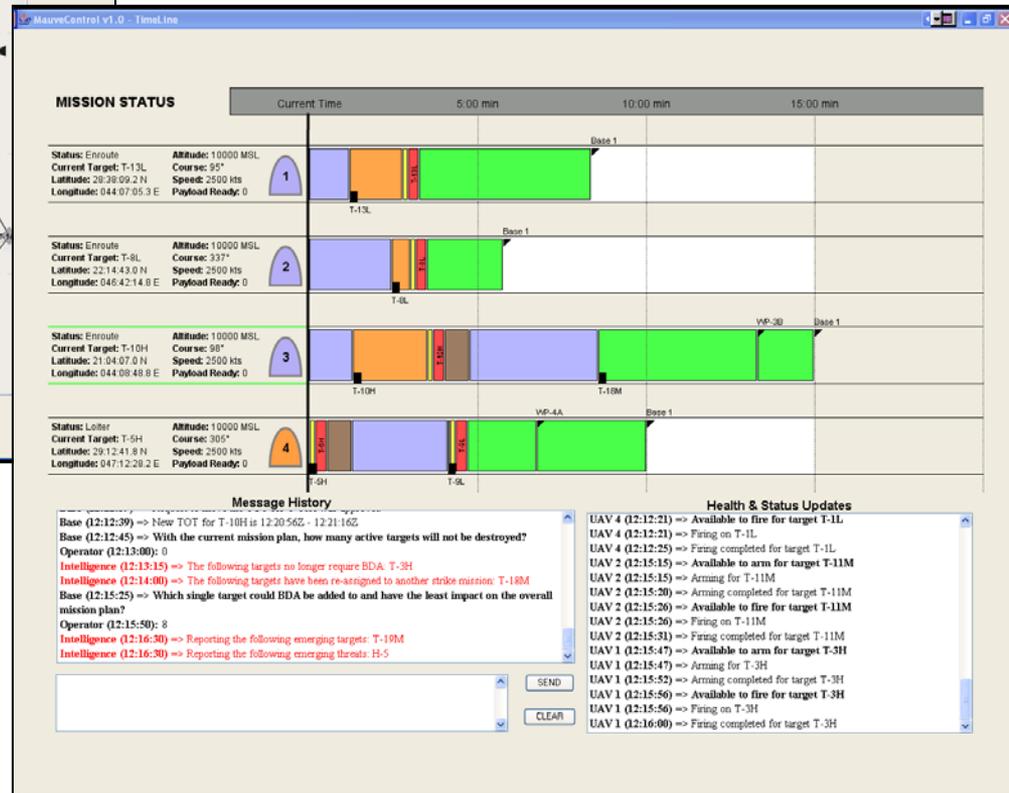
12:15:24



Results and Discussion – Example Test Session – 9



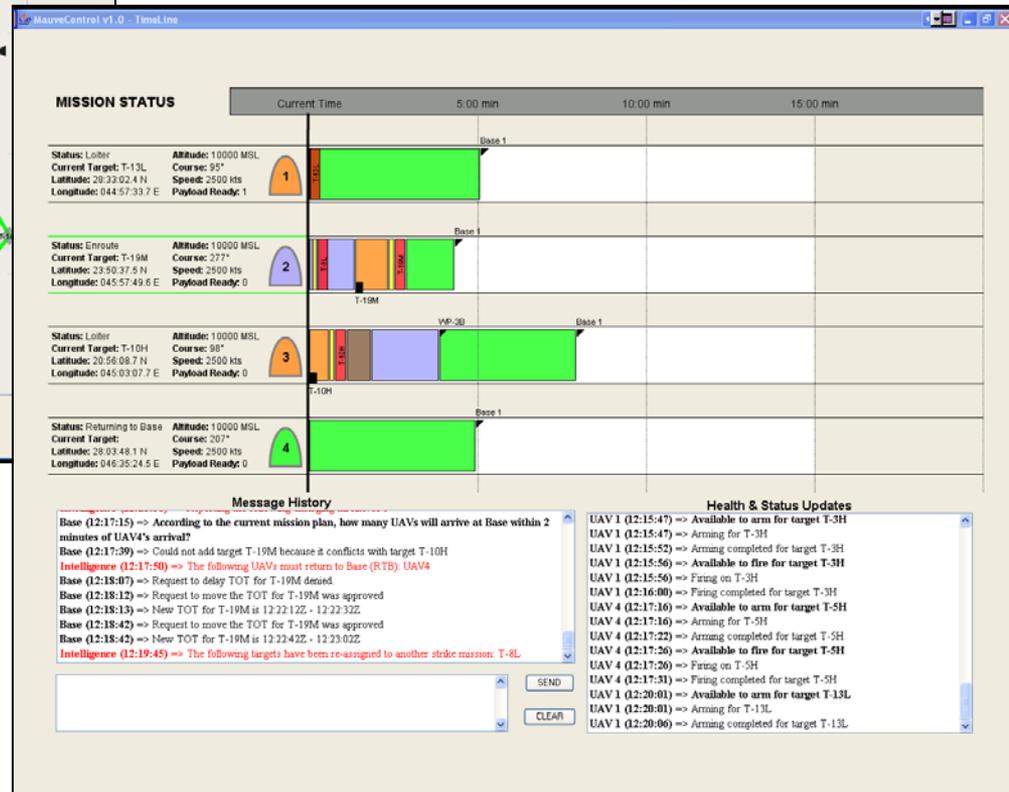
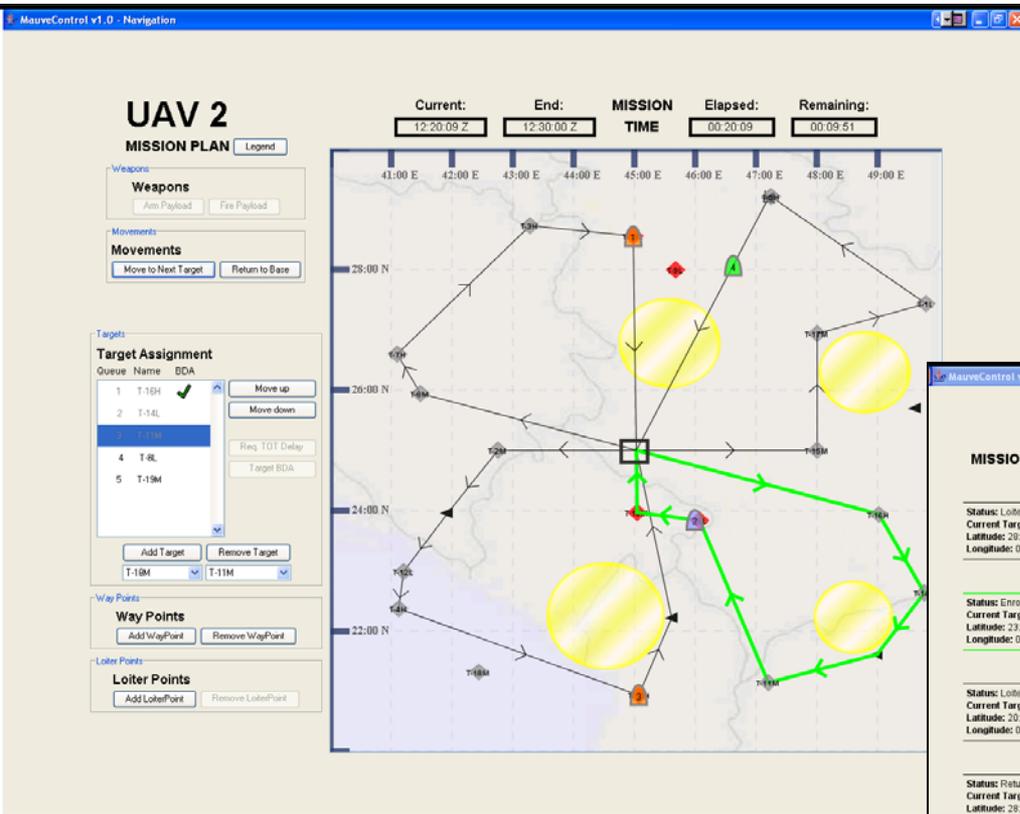
Super-Active Automation High Re-planning Scenario



12:17:14



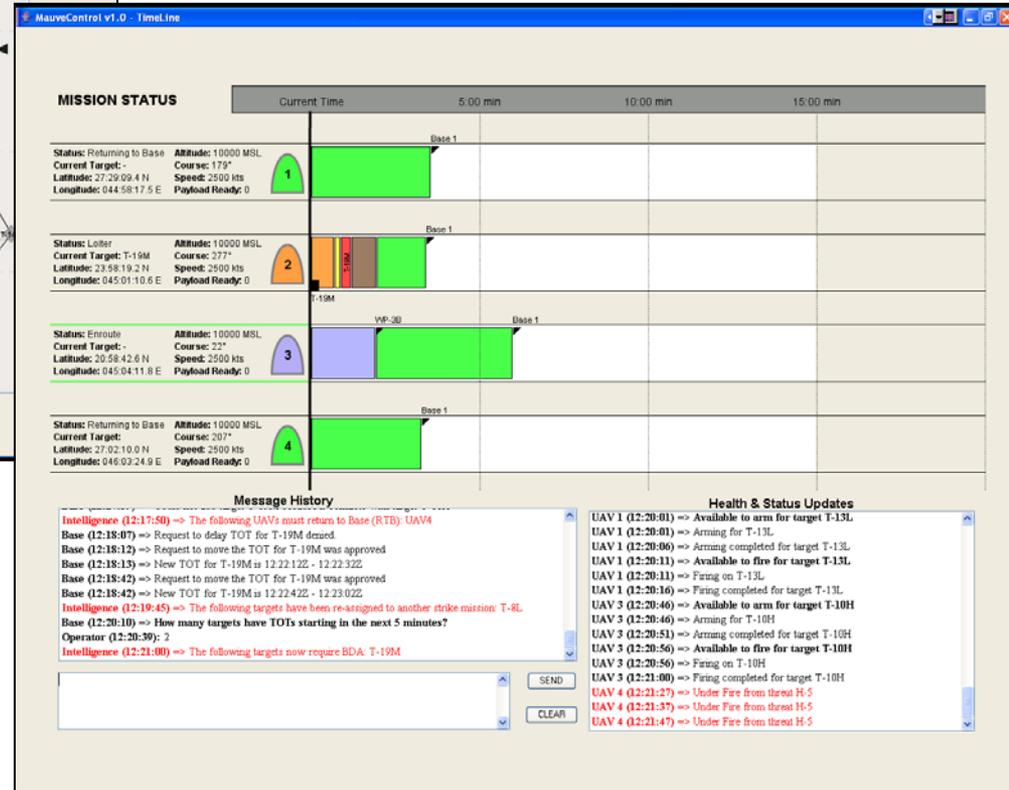
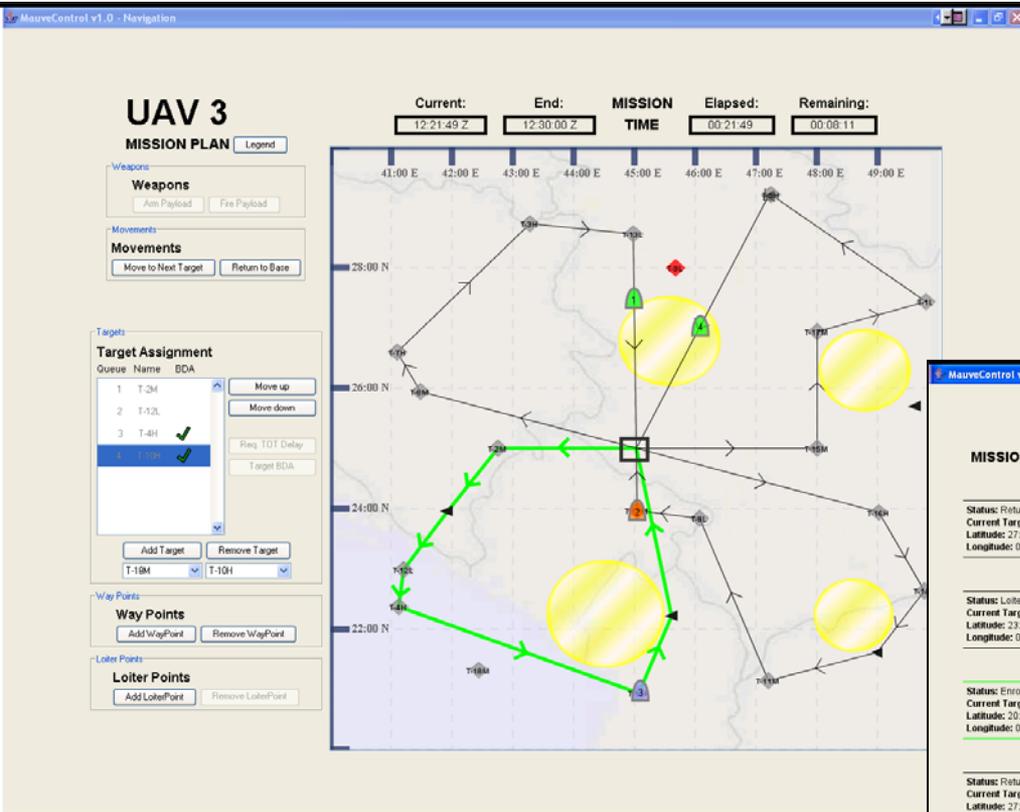
Super-Active Automation High Re-planning Scenario



12:20:09



Super-Active Automation High Re-planning Scenario

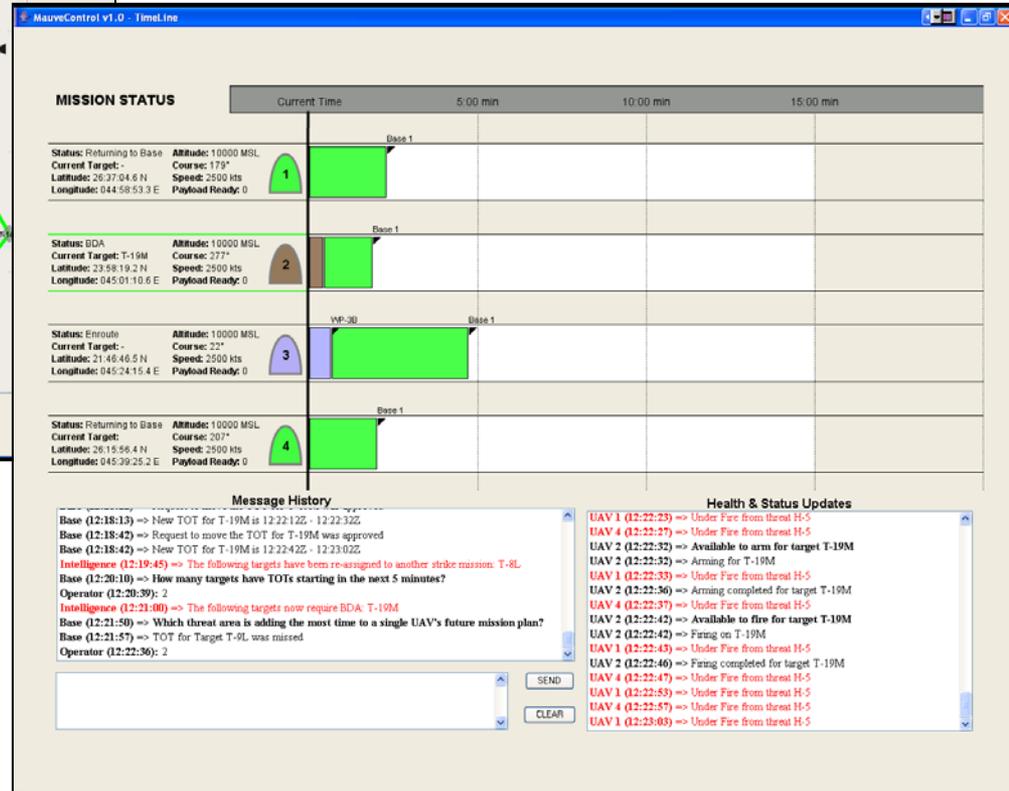
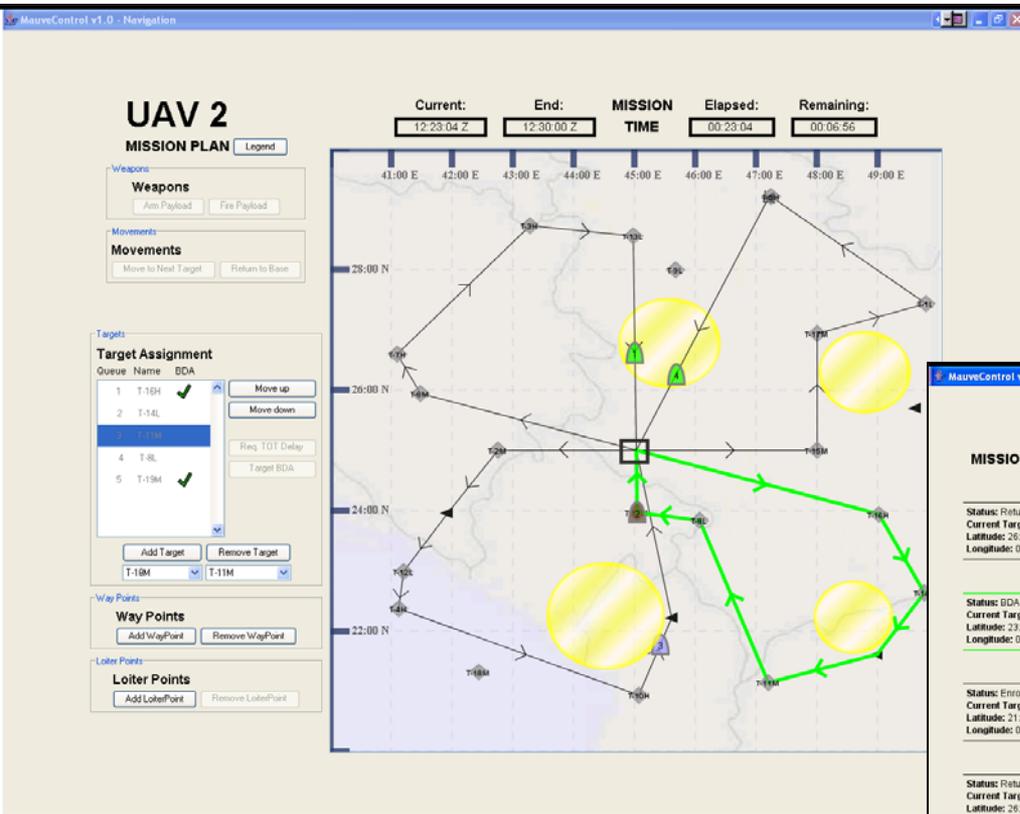


12:21:49



Results and Discussion – Example Test Session – 12

Super-Active Automation High Re-planning Scenario



12:23:04

