Data Mining Techniques Applied to Urban Terrain Command & Control Experimentation

Janet O'May
jomay@arl.army.mil
(410) 278-4998

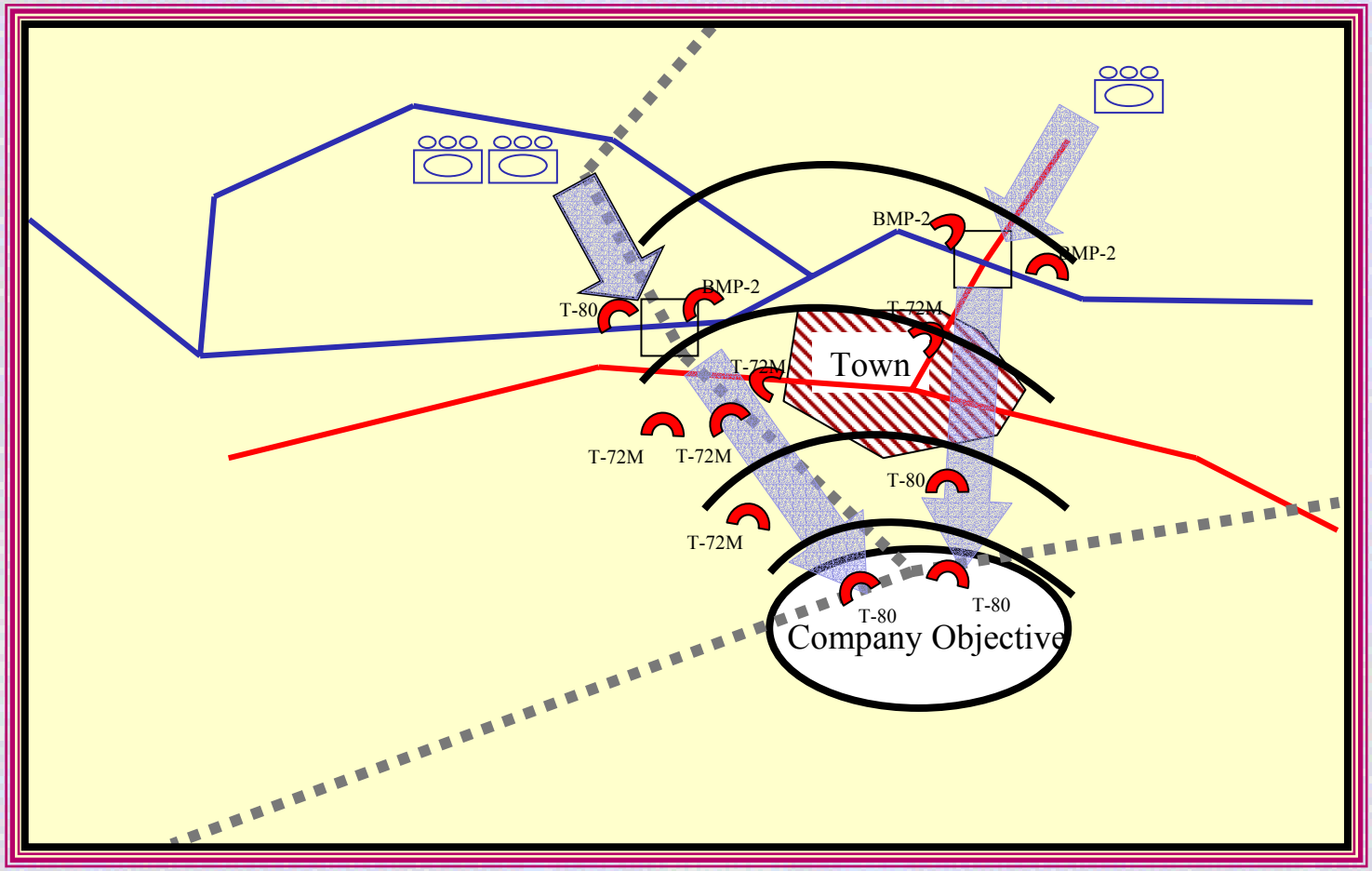
Co-Authors:
Barry Bodt and Eric Heilman

Computational and Information Sciences Directorate
Army Research Laboratory (ARL)
The U.S. Army's Corporate Laboratory

Previous Work

- Simulated Southwest Asia scenario using One Semi-Automated Forces (OneSAF) Testbed Baseline (OTB) for 228 plays of the same battle
- Used a Killer/Victim Scoreboard (KVS) to collect metrics [143 metrics per three time slice]
- Used statistical data mining approaches to relate battle outcome to metrics
- Concluded: great potential for identifying key metrics in the battle worth tracking and/or for suggesting course of action changes

Scenario


BMP-2
BMP-2
T-80
BMP-2
T-72M
T-72M
Town
T-72M
T-72M
T-80
T-72M
T-80
T-80
Company Objective

OneSAF Modification
Killer/Victim Scoreboard

- Firer and Target Identity and Location
- Type of Ammo
- Range
- Outcome

```
Time Stamp 1010070890
Vehicle ID 1076
Firer ID 1087
Projectile 1143670848
Firer Position:  X = 220217.00  Y = 146765.00  Z = 12.37
Target Position:  X = 222454.38  Y = 149117.80  Z = 9.99
Vehicle 1076: Hit with 1 "munition_USSR_Spandrel" (0x442b0840)
  Comp DFDAM_EXPOSURE_HULL, angle 19.53 deg Disp 0.889701 ft
  Kill Thermometer is: Pk:1.00, Pmf:1.00, Pf:0.90, Pm:0.80 Pn:0.80
RANGE  3246.773576
r = 0.990835 kill_type = MF
```

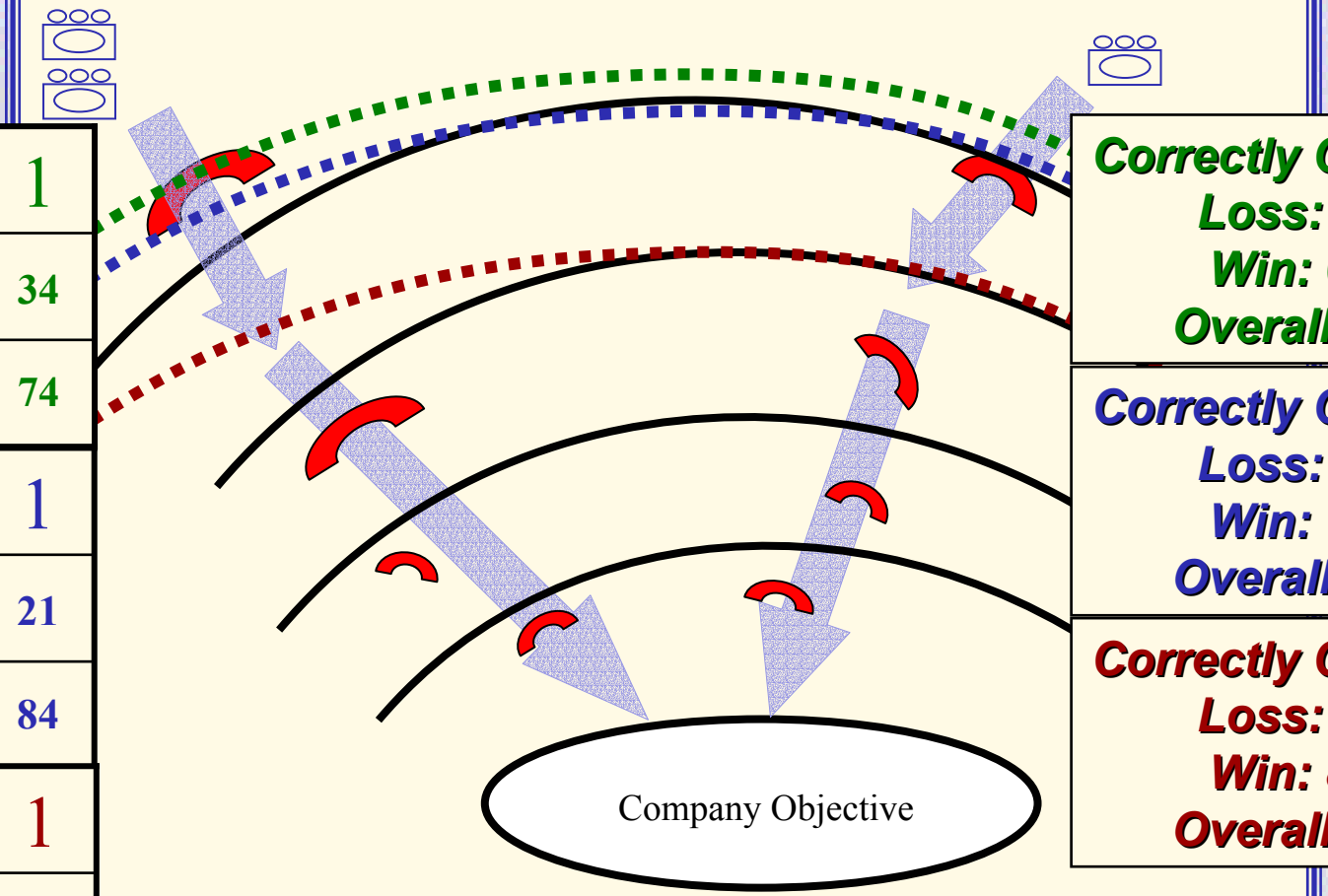
```
1076 100A41 vehicle_US_M1
1087 100A23 vehicle_USSR_BMP2
```

Analysis

Pred / Obs	0	1
0	85	34
1	35	74

Pred / Obs	0	1
0	98	21
1	25	84

Pred / Obs	0	1
0	105	14
1	20	89



Correctly Classified
Loss: 71%
Win: 67%
Overall: 70%

Correctly Classified
Loss: 82%
Win: 77%
Overall: 80%

Correctly Classified
Loss: 88%
Win: 82%
Overall: 85%

Slice 1 ~ 2000m Or ~ 5 ½ minutes

Slice 2 ~ 4000m Or ~ 10 minutes

Slice 3 ~ 5800m Or ~ 20 minutes

Method Comparison

Percent Correct Classification by Stopping Time and Method

Stopping Time (min)	Discriminant Analysis	CART	Logistic Regression
5 ½	70%	70%	69%
10	80%	75%	74%
20	85%	82%	85%

Current Experiment

- Change terrain to urban
- Involve Dismounted Infantry (DI)
- Use Dismounted Infantry Semi-Automated Forces (DISAF) Simulation Software
- Develop urban scenario

DISAF Challenges

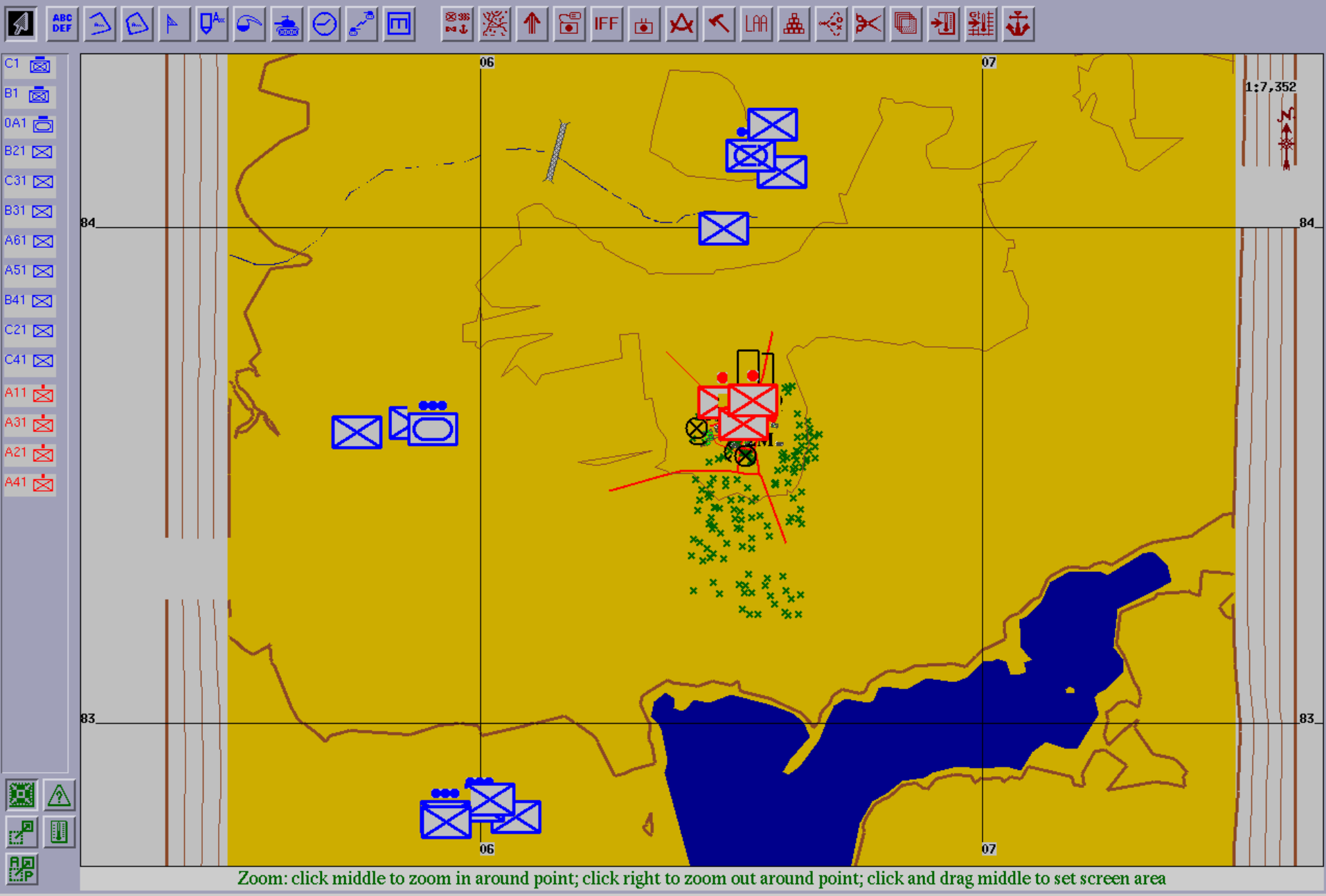
- Compiling – need an older version of GNU C and C++ (version 2.91.66)
- KVS Code developed for OneSAF at ARL did not easily insert into DISAF
- Fireteams tend to move better when tasked as individual rather than as a team
- However this breaks down for the “clear room” task which requires a full fireteam
- Vehicles tend to not enter the city sector
- DI entities at times get stuck in buildings and tunnels
- If entities can not determine the proper route they go to the bottom of the terrain

Current Urban Scenario

Location: City sector based on the McKenna MOUT (Military Operations on Urbanized Terrain) site

Scenario: Attack Phase I

- Isolate area, three-pronged encirclement to reduce threat forces from perimeter
- Carried out by 2 M2s from the North and 2 M2s from the Southwest and a headquarters (HQ) attachment of 2 M1A1s and 1 M2 entering from the West
- Initial resistance from 3 BMPs and 2 T-80s around the perimeter, a 3rd T-80 is in the center flanking the objective



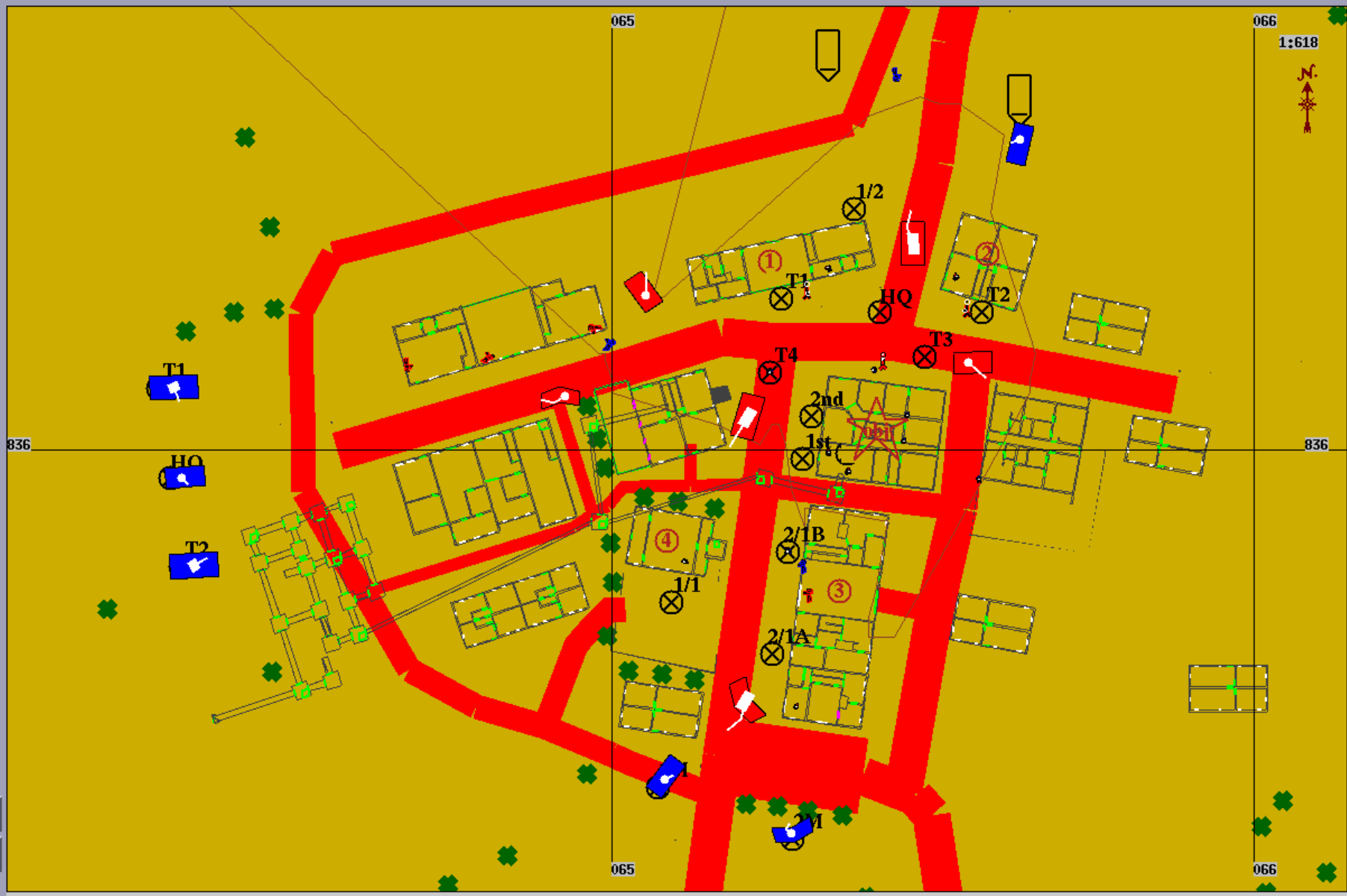
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area

Select an item to edit


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
C1 B1 0A1 B21 C31 B31 A61 A51 B41 C21 C41 A11 A31 A21 A41
1:485
T1 HQ T2 T3 T4 2nd 1st 1/2 2/1B 1/1 2/1A 1M
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit

OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001

File Scale Features Map Scale Show As Special Local Designation HHours Privilege 15:05:44 GMT



Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area

Select an item to edit

Current Urban Scenario

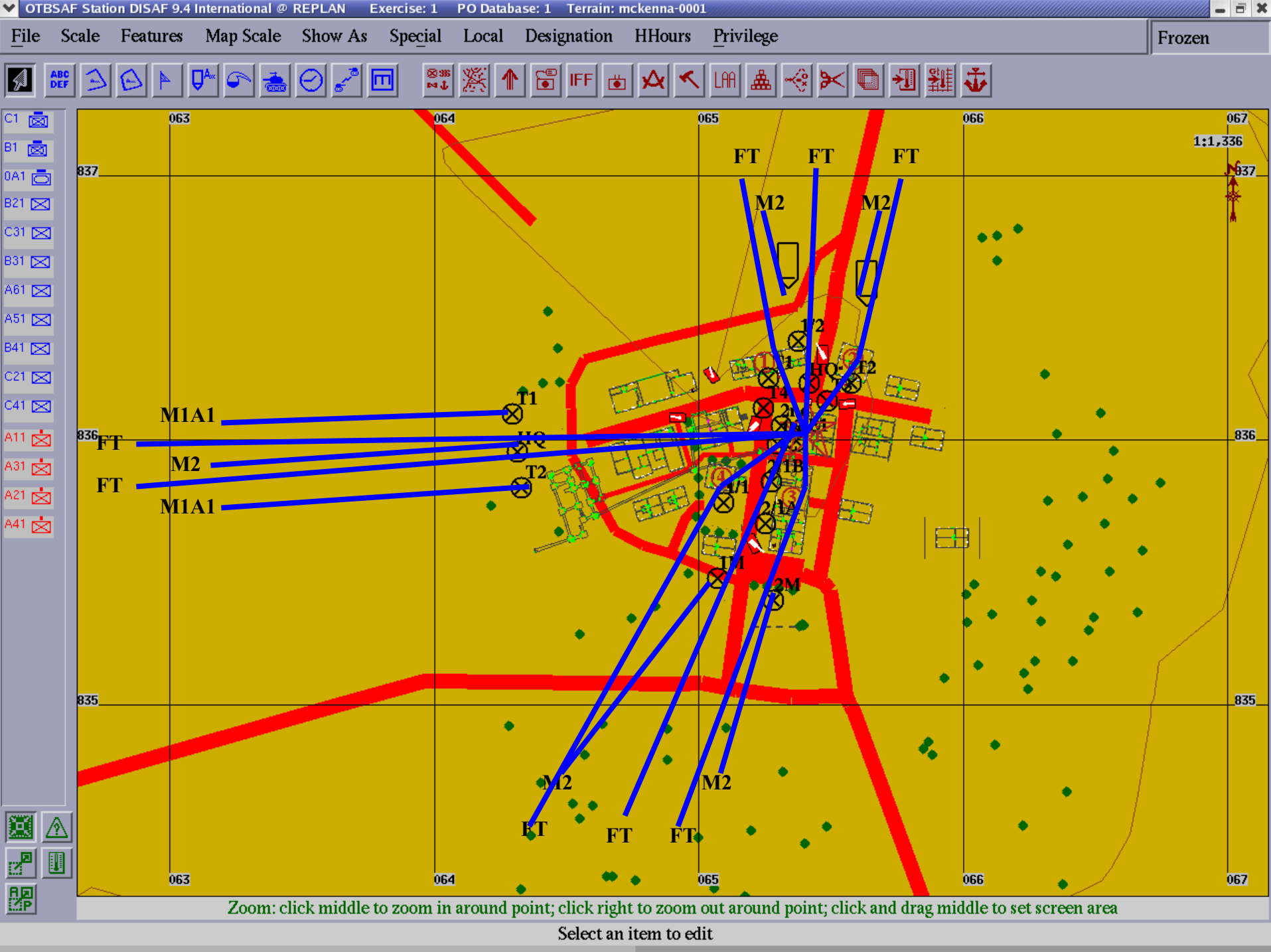
Scenario: Attack Phase 2

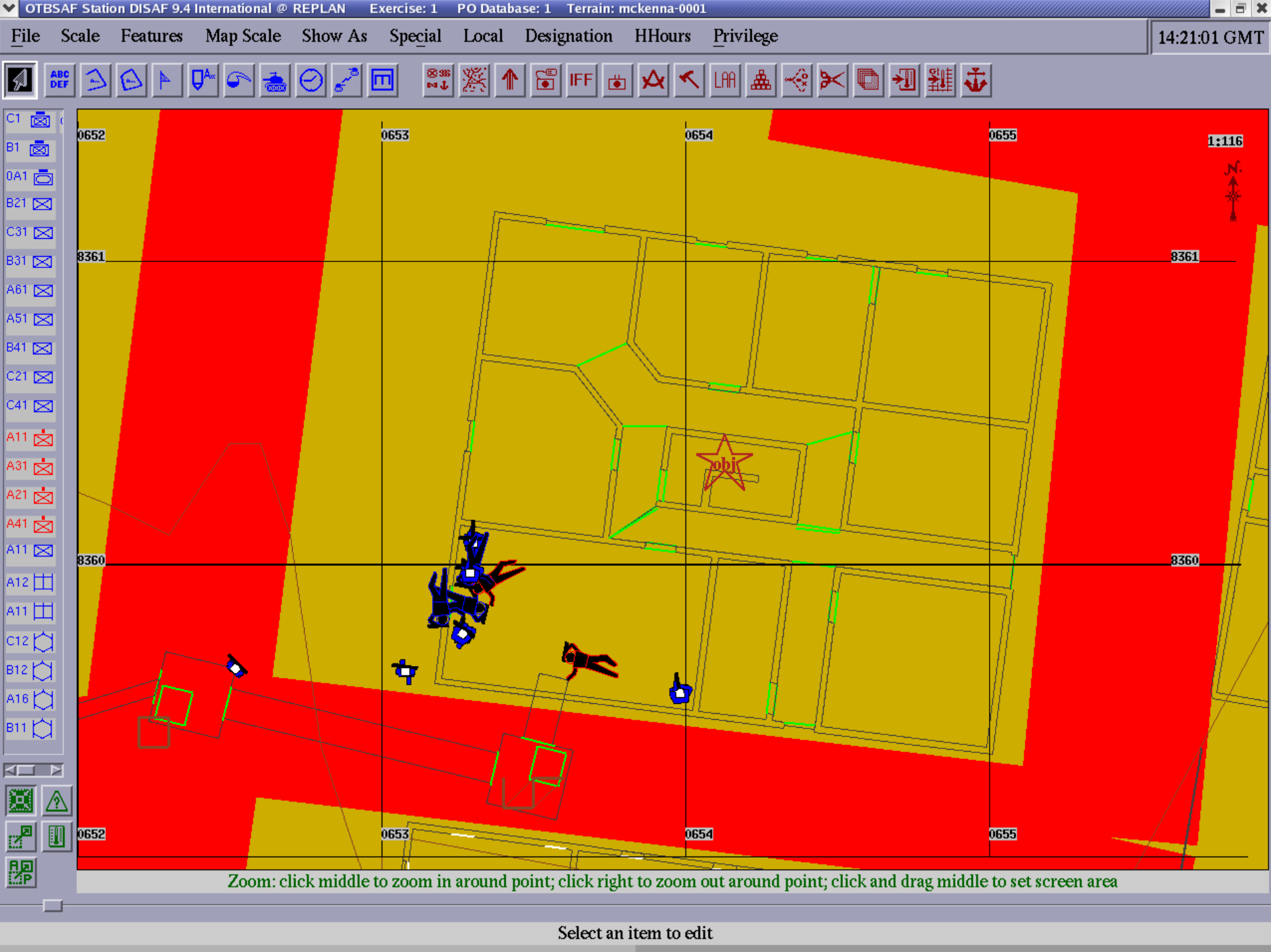
- Eight fireteams (FT) enter sector behind armored vehicles
- Carried out by 3 FTs from both the North and from the Southwest
- Northern teams clear separate buildings (1 & 2) and continue on to secure objective
- Southwestern teams clear separate buildings (3 & 4) and continue to objective
- Two Western FT (HQ) proceed directly to the objective

Current Urban Scenario

Scenario: Attack Phase 2

- Interior resistance provided by opposition DI in the five critical buildings and also in a key vantage point building on the Northwest side of the sector
- Three additional opposition DI stationed outside buildings 1, 2, and the objective


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
Frozen
1:1,336
063 064 065 066 067
837 836 835
FT FT FT
M2 M2
M1A1
FT
M2
FT
M1A1
T1 HQ T2
M2 M2
FT FT FT
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
14:21:01 GMT
1:116
0652
0653
0654
0655
8361
8360
obj
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit

Simulation Data

Predictors

- 444 variables, *but only 75 runs so far*
- Two time slices (372 seconds and 480 seconds)
- Hits taken by Blue and hits by Blue fire involving all relevant vehicles, fireteams, and buildings
- Status of all entities

Responses

- Taking the objective (1 or 0)
- Establishing a foothold in the city (0, 1, 2, 4)
- MOUTscore (0 to 8 with buildings under control with minimum casualties)

Potential Analytical Methods

- Discriminant Analysis
- Cart
- Logistic Regression
- Multiple Regression
- Neural Networks

•Dr. Barry Bodt babodt@arl.army.mil

A Discriminant Model

Data: Discriminant Function Analysis Summary (FINALMOUTEX1b)*

Discriminant Function Analysis Summary (FINALMOUTEX1b)
Step 7, N of vars in model: 7; Grouping: FootHold (4 grps)
Wilks' Lambda: .32024 approx. F (21,187)=4.3384 p< .0000

N=75	Wilks' Lambda	Partial Lambda	F-remove (3,65)	p-level	Toler.	1-Toler. (R-Sqr.)
F8TS2	0.488062	0.656143	11.35459	0.000004	0.910182	0.089818
A23TS2	0.380200	0.842291	4.05682	0.010497	0.921236	0.078764
RTA11TS2	0.358180	0.894073	2.56701	0.062007	0.685511	0.314489
BMC12TS2	0.374920	0.854151	3.69965	0.016001	0.877811	0.122189
F5TS2	0.354554	0.903217	2.32167	0.083316	0.967066	0.032934
M3T13FS2	0.344559	0.929415	1.64548	0.187509	0.919428	0.080572
M2A21TS2	0.341595	0.937480	1.44494	0.237843	0.720540	0.279461

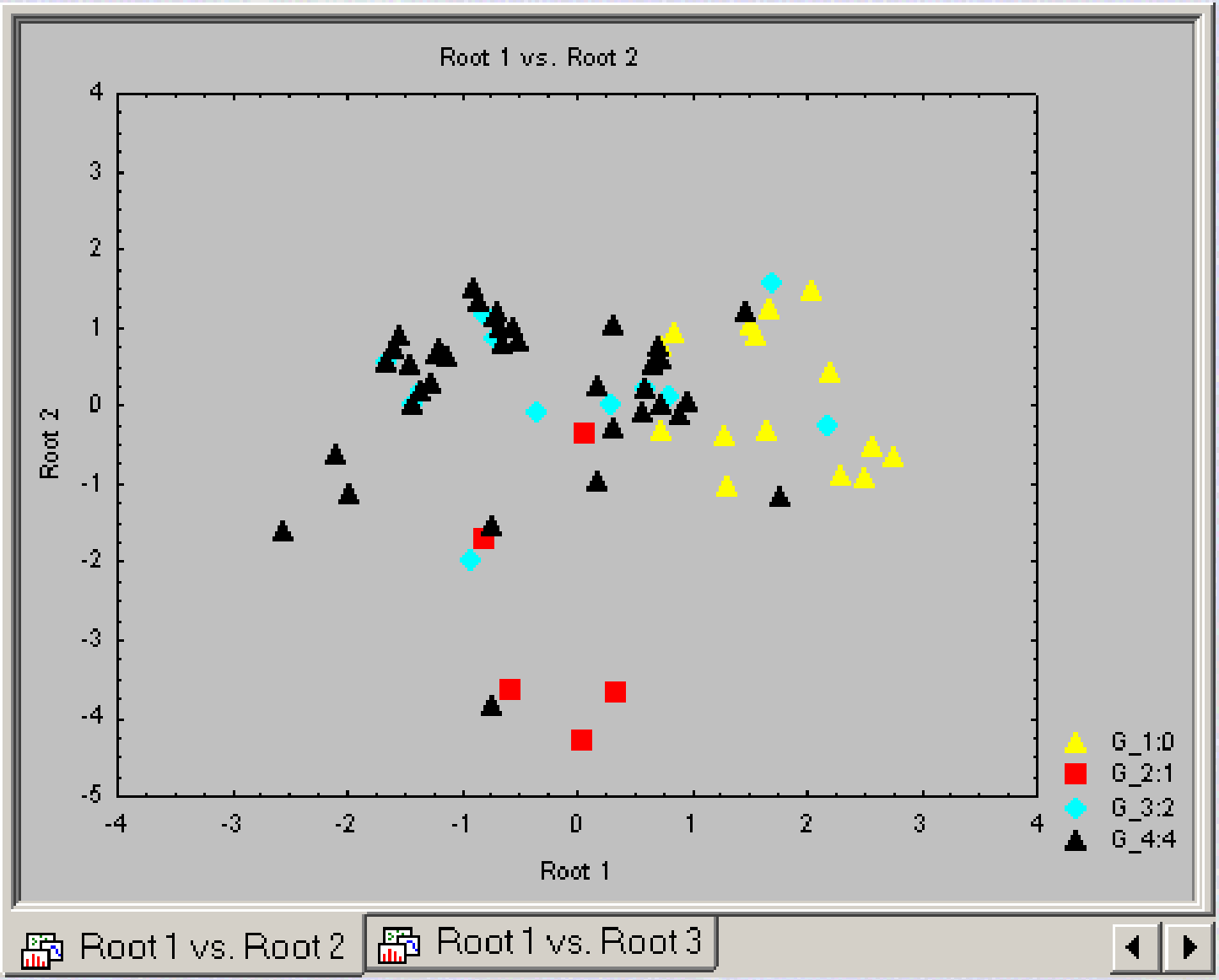
Root Means

Means of Canonical Variables (FI

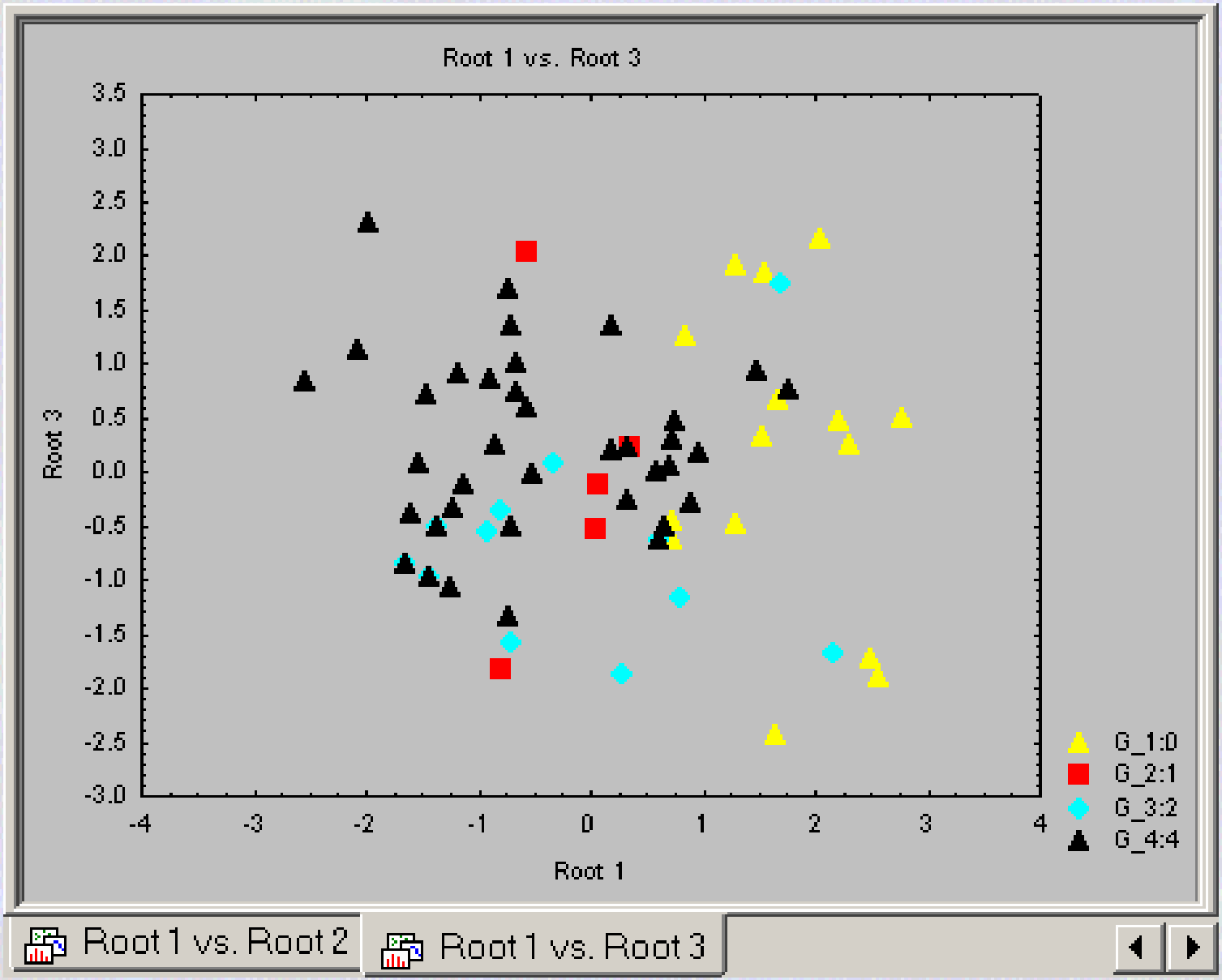
Group	Root 1	Root 2	Root 3
G_1:0	1.645990	0.11182	0.072312
G_2:1	-0.182453	-2.73915	-0.044188
G_3:2	-0.273964	0.26408	-0.658511
G_4:4	-0.539702	0.20524	0.207078

Means of Canonical Variables (FINALMOUTE

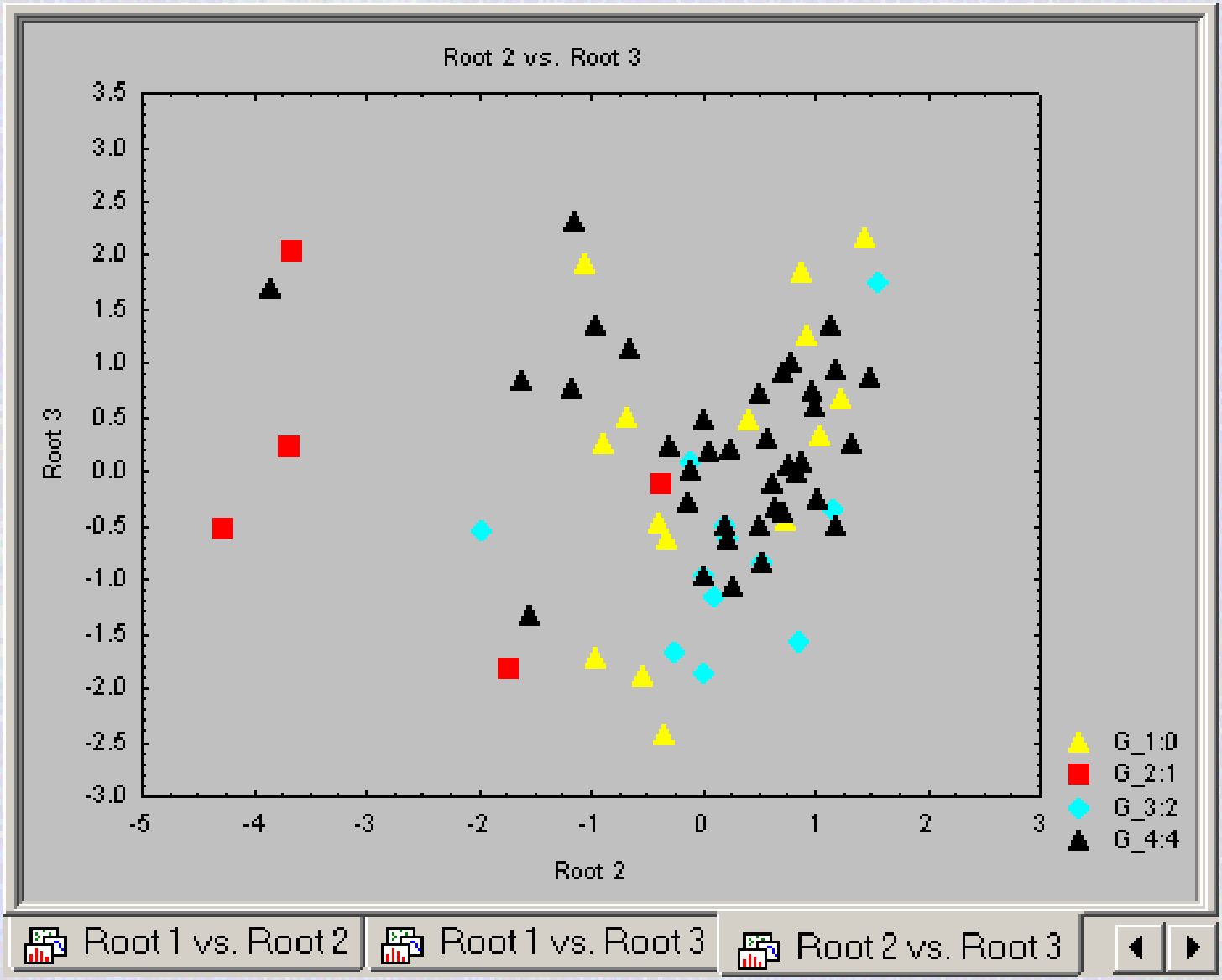
Group Separation



Group Separation



Group Separation



Classification Efficiency

Classification Matrix (FINALMOUTEX1b)
Rows: Observed classifications
Columns: Predicted classifications

Group	Percent Correct	G_1:0 p=.21333	G_2:1 p=.06667	G_3:2 p=.18667	G_4:4 p=.53333
G_1:0	75.00000	12	0	0	4
G_2:1	60.00000	0	3	1	1
G_3:2	21.42857	2	0	3	9
G_4:4	92.50000	2	1	0	37
Total	73.33334	16	4	4	51

Classification Matrix (FINALMOUTEX1b)

Metric Description

F8TS2-Status of FT 8 (B4) at TS2

F5TS2-Status of FT 5 (A6) at TS2

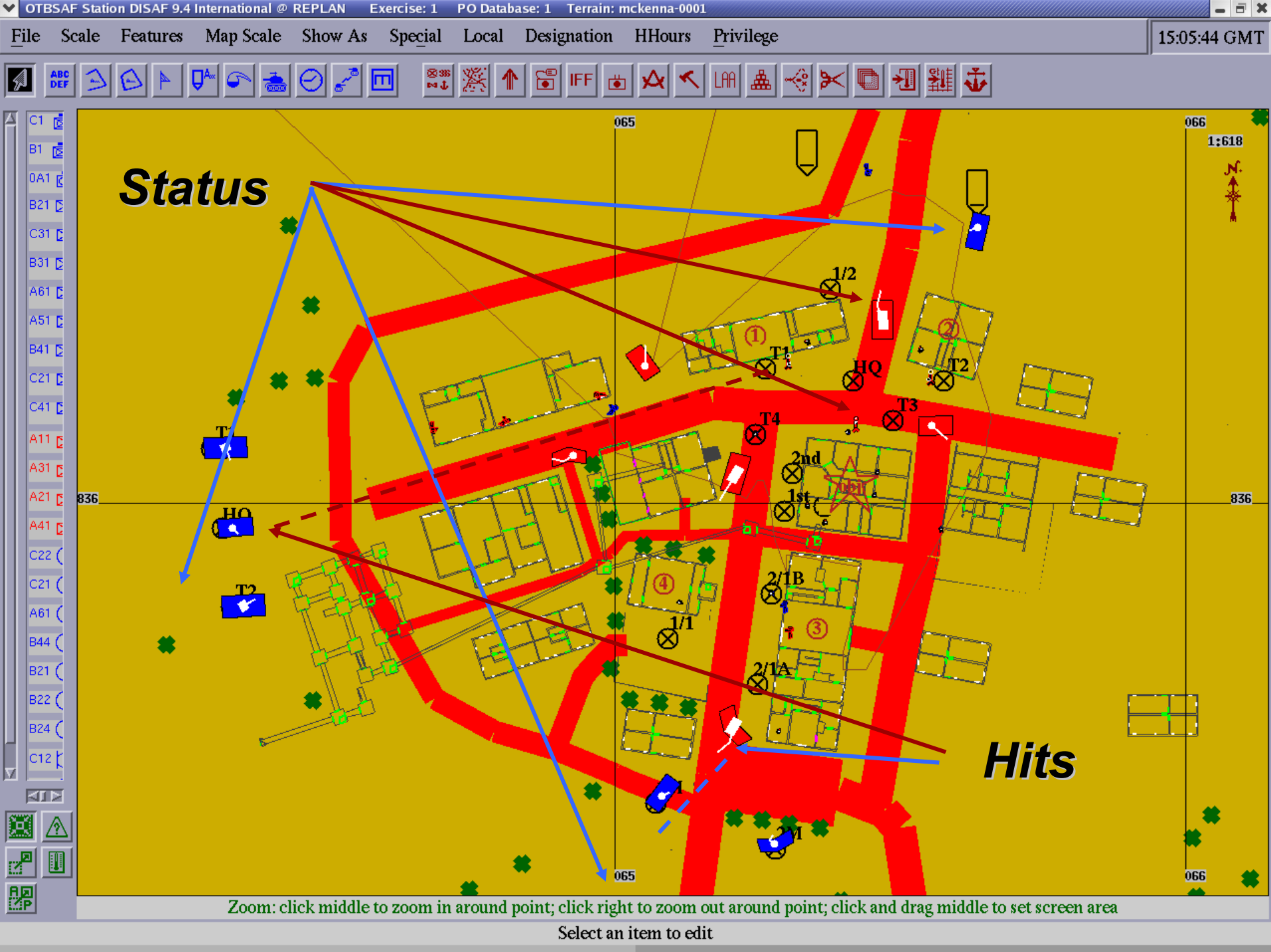
M3T13FS2-Hits by M2-3 (B11, B12) at A13-T80 at TS2

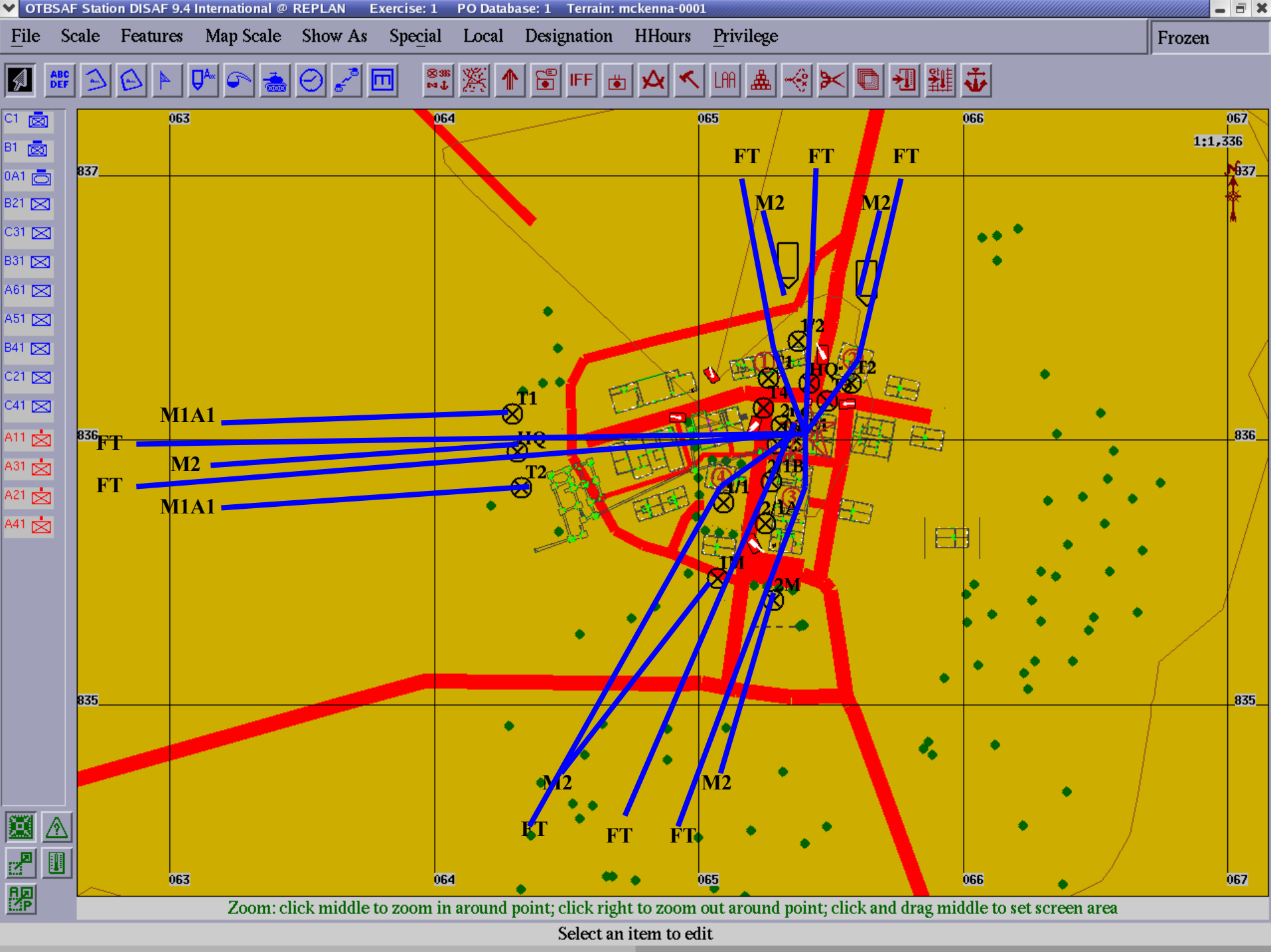
A23TS2-Status of A23-DI at TS2

BMC12TS2-Status of M2 C12 at TS2

RTA11TS2-Status of T80 A11 at TS2

M2A21TS2-Hits taken by M2-2 (A16) by A21-DI at TS2


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
15:05:44 GMT
Status
Hits
065
066
836
1:618
1/2
T1
HQ
T2
T3
T4
2nd
1st
2/1B
1/1
2/1A
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
Frozen
1:1,336
063 064 065 066 067
837 836 835
FT FT FT
M2 M2
M1A1
FT
M2
FT
M1A1
T1 HQ T2
M2 M2
FT FT FT
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit

Metric Interpretation

FT8 important because mission took it first to Bldg 3 and positioning kept it away from being targeted by Bldg 5

FT5 important because after passing Bldg 5, would move South away from Red DI in approach to objective

M3T13FS2-Hits on T80 – Southwest approach

A23TS2-Status of A23-DI – outside Building 2

BMC12TS2-Status of M2 - Success of North approach

RTA11TS2-Status of T80 on North

M2A21TS2-Hits taken by M2-2 – West approach

Acknowledgements

- Brian Comer of PEO STRI
- Daryl Siddon of SAIC

Without their help, we would never have compiled DISAF!

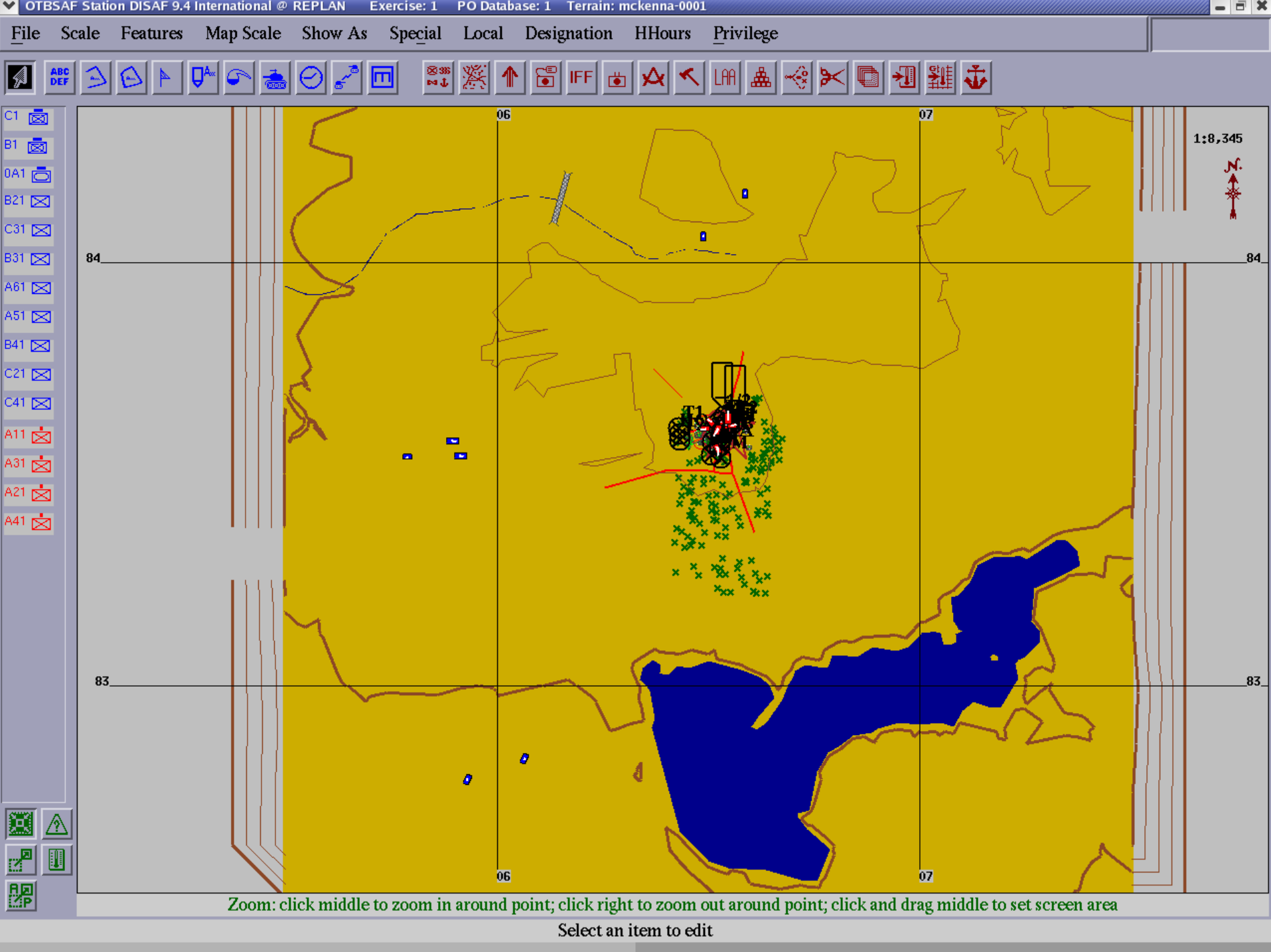
Directions/Conclusion

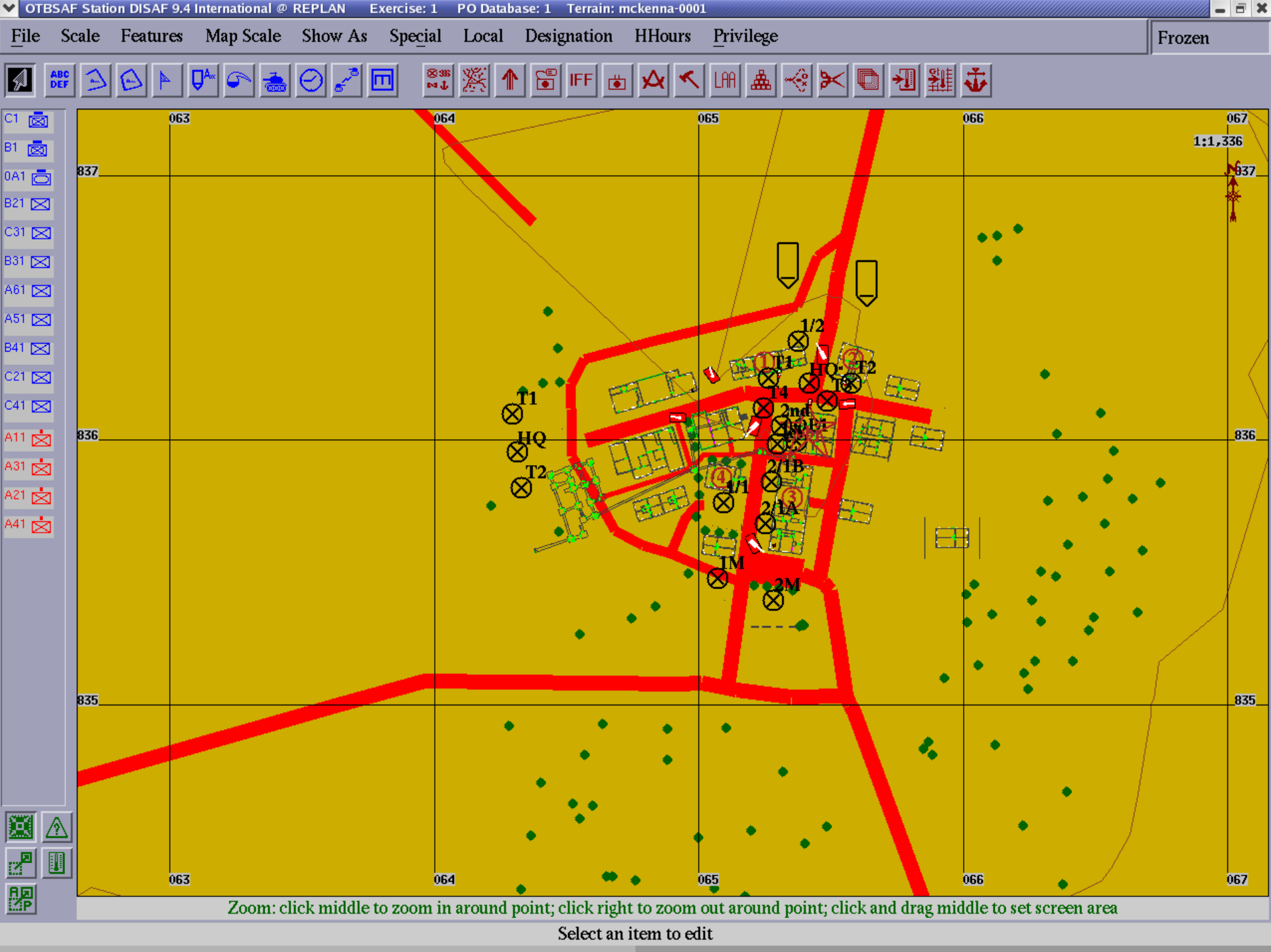
- More involved and sophisticated opposition force
- Changes to KVS
- Overall improved scenario
- Establish data from more time periods to provide information on battle progression
- Collect information from a larger number of battles
- Use a greater variety of statistical tools, to include work in the microarray arena
- Data mining combat simulations holds great promise for understanding battles if one believes the simulations and statistical methods will continue to improve.

Questions ???

Richard C. Kaste
Team Leader
rck@arl.army.mil
(410) 278-7781

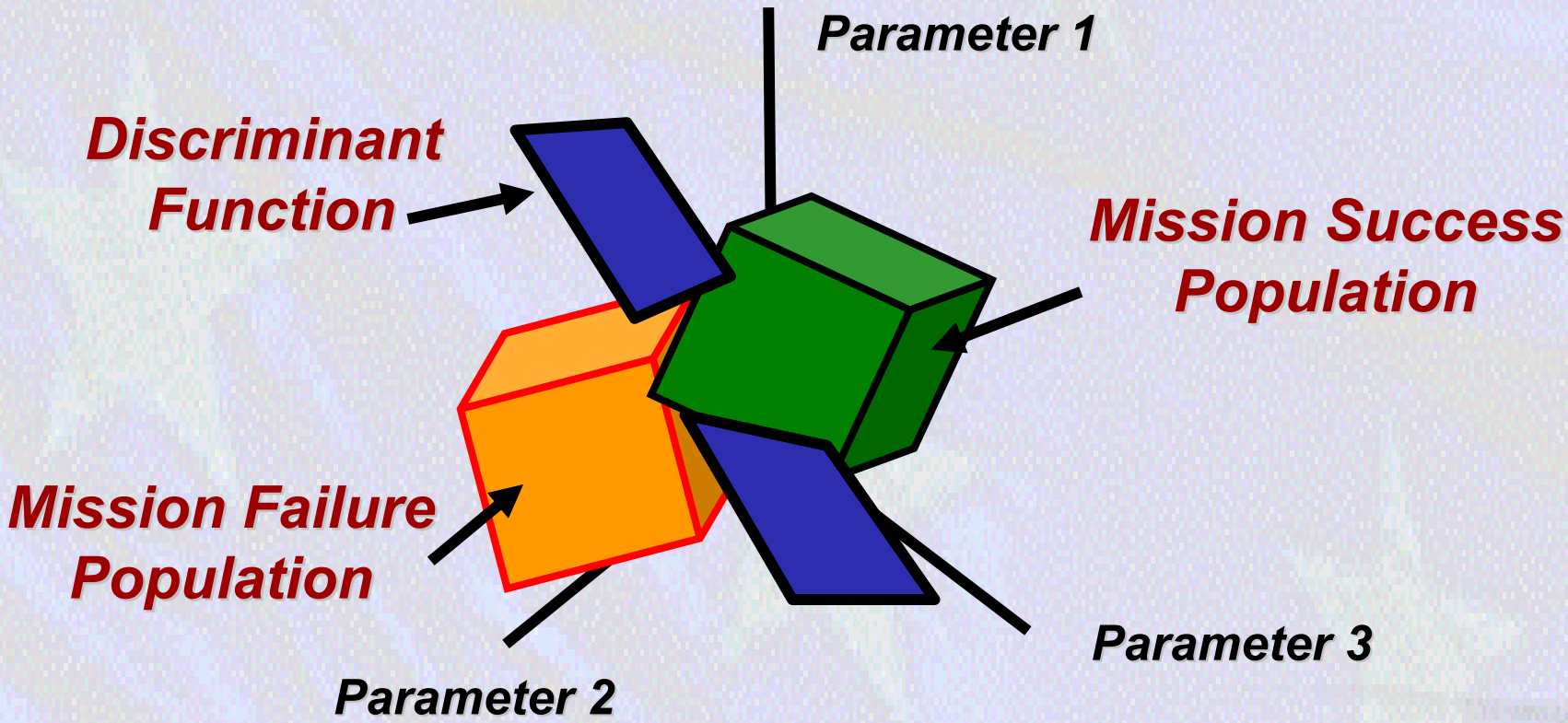
Janet O'May
jomay@arl.army.mil
(410) 278-4958


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
C1
B1
0A1
B21
C31
B31
A61
A51
B41
C21
C41
A11
A31
A21
A41
06
07
84
83
1:8,345
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit


OTBSAF Station DISAF 9.4 International @ REPLAN Exercise: 1 PO Database: 1 Terrain: mckenna-0001
File Scale Features Map Scale Show As Special Local Designation HHours Privilege
Frozen
1:1,336
063 064 065 066 067
837 836 835
1/2
T1
HQ
T2
T4
2nd
2/1B
1/1
2/1A
1M
2M
Zoom: click middle to zoom in around point; click right to zoom out around point; click and drag middle to set screen area
Select an item to edit

Discriminant Analysis

- **Maximizes** $|a'(\bar{x}_1 - \bar{x}_2)| \; s.t. \; a'Sa = 1$
- **Assumes multivariate normal predictors with common covariance matrix Σ but different mean vectors μ_1 and μ_2**



Standardized Coefficients

Variable	Standardized Coefficients (FINALM for Canonical Variables		
	Root 1	Root 2	Root 3
F8TS2	-0.916437	0.127259	-0.088794
A23TS2	0.345706	0.556590	-0.246588
RTA11TS2	0.533809	0.179519	-0.411562
BMC12TS2	0.266831	-0.604834	-0.143915
F5TS2	-0.398093	0.267274	0.203862
M3T13FS2	-0.088558	-0.259234	-0.699959
M2A21TS2	-0.059910	0.218665	0.824142
Eigenval	0.791787	0.568677	0.110980
Cum.Prop	0.538102	0.924577	1.000000

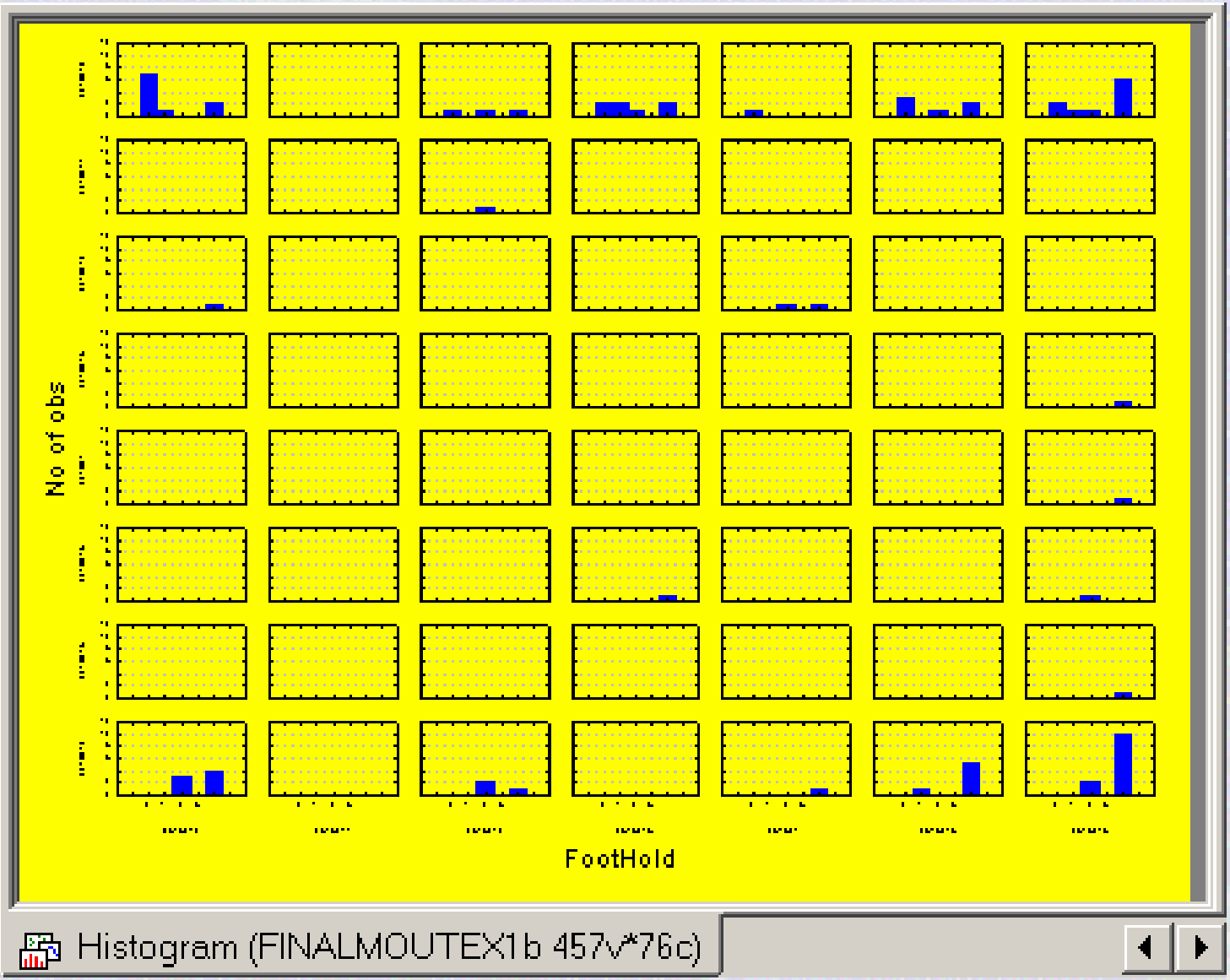
Standardized Coefficients (FINALMOUTEX1b)

Testing Roots

Roots Removed	Chi-Square Tests with Successive Roots Removed (FINA					
	Eigen-value	Canonicl R	Wilks' Lambda	Chi-Sqr.	df	p-level
0	0.791787	0.664754	0.320239	78.00018	21	0.000000
1	0.568677	0.602097	0.573799	38.05006	12	0.000151
2	0.110980	0.316060	0.900106	7.20912	5	0.205546

Chi-Square Tests with Successive Roots Removed (FINALMOUTEX1b)

FT3 (Y) by FT7 (X) by Foothold Frequency



FT5 (Y) by FT8 (X) by Foothold Frequency