APPENDIX C Validation & Confidence Building

Emerging-State Actor Model (E-SAM)

**Appendix C Table of Contents**

[C-1 Introduction 3](#_Toc507503117)

[C-2 Boundary Adequacy 4](#_Toc507503118)

[Boundary Tests 4](#_Toc507503119)

[C-3: Structure Assessment 14](#_Toc507503120)

[C-4: Dimensional Consistency 39](#_Toc507503121)

[C-5: Parameter Assessment 39](#_Toc507503122)

[C-6: Extreme Condition 40](#_Toc507503123)

[C-7: Integration Error 53](#_Toc507503124)

[C-8: Behavior Reproduction 58](#_Toc507503125)

[C-9: Behavior Anomaly 61](#_Toc507503126)

[C-10: Family Member Test 64](#_Toc507503127)

[C-11: Surprise Behavior 90](#_Toc507503128)

[C-12: Sensitivity Analysis 91](#_Toc507503129)

[C-13: System Improvement 106](#_Toc507503130)

**TABLE OF APPENDIX FIGURES**

[Figure 1: Time Boundary Test on Territory 20yr 5](#_Toc507503131)

[Figure 2: Time Boundary Test on Population 20yr 6](#_Toc507503132)

[Figure 3: Time Boundary Test with Demographics Activated 7](#_Toc507503133)

[Figure 4: Interventions Boundary Test on Total Population 8](#_Toc507503134)

[Figure 5: Intervention Boundary Test on Territory 9](#_Toc507503135)

[Figure 6: Intervention Boundary Test External Intervention Sizes 10](#_Toc507503136)

[Figure 7: Intervention Boundary Test - Local Intervention Sizes 11](#_Toc507503137)

[Figure 8: Intervention Boundary Test - Average Combatant Experience 12](#_Toc507503138)

[Figure 9: Intervention Boundary Test - Offensive Stance of Green 13](#_Toc507503139)

[Figure 10: Structure Assessment - Conservation of Mass Error 15](#_Toc507503140)

[Figure 11: Structure Assessment - Conservation of Mass Correction 15](#_Toc507503141)

[Figure 12: Structure Assessment - Ethno by Actor Sufficiency Values 16](#_Toc507503142)

[Figure 13: Structure Assessment - Free Lunch Error Territory 17](#_Toc507503143)

[Figure 14: Structure Assessment - Free Lunch Error Total Combatants 17](#_Toc507503144)

[Figure 15: Structure Assessment - Free Lunch Error Total Terrorist Attacks 18](#_Toc507503145)

[Figure 16: Structure Assessment - Source of Free Lunch Error 19](#_Toc507503146)

[Figure 17: Fixing Free Lunch Errors - Sending Funds Abroad & Spending Prioritization 20](#_Toc507503147)

[Figure 18: Fixing Free Lunch Errors - Allocation of Essential & Non-Essential Budget Priorites to Payroll and Detainee Benefits 21](#_Toc507503148)

[Figure 19: Fixing Free Lunch Errors - Payroll Gap Structure added to Total Defections Rate 21](#_Toc507503149)

[Figure 20: Fixing Free Lunch Errors - Detention Benefits Structure added to Defections within Prison 22](#_Toc507503150)

[Figure 21: Fixing Free Lunch Errors - Gaps based on prioritization of Spending 22](#_Toc507503151)

[Figure 22: Fixing Free Lunch Errors - Comparison of Total Defections Rates 23](#_Toc507503152)

[Figure 23: Fixing Free Lunch Errors - Comparison of Detainees in Prison 24](#_Toc507503153)

[Figure 24: Structure Assessment - Information Error 25](#_Toc507503154)

[Figure 25: Structure Assessment - Correction to Information Error in Perception of Momentum 26](#_Toc507503155)

[Figure 26: Structure Assessment - Allocation of Conventional Forces Historical 27](#_Toc507503156)

[Figure 27: Structure Assessment - Allocation of Conventional Forces without Intervention 28](#_Toc507503157)

[Figure 28: Structure Assessment - Correction of Information Error in Ethnographic Perception of Momentum 29](#_Toc507503158)

[Figure 29: Structure Assessment - Relative Ethnographic Perception of Momentum Historical 30](#_Toc507503159)

[Figure 30: Structure Assessment - Ethnographic Perception of Relative Momentum without Intervention 31](#_Toc507503160)

[Figure 31: Unit Consistency 39](#_Toc507503161)

[Figure 32: Extreme Condition - Test Structure 40](#_Toc507503162)

[Figure 33: Integration Test - Differences in KIA per M by Method 56](#_Toc507503163)

[Figure 34: Integration Test - Differences in Local Opposition to Actor by Method 56](#_Toc507503164)

[Figure 35: Integration Test - Integration Error Visible under Certain Conditions 57](#_Toc507503165)

[Figure 36: Integration Test - Structural Fix to Integration Error Demonstrated 58](#_Toc507503166)

[Figure 37: Loop Knockout of Revenue Feedback effect on Territory Controlled 63](#_Toc507503167)

[Figure 38: Loop Knockout of Revenue Feedback effect on Total Combatants 64](#_Toc507503168)

[Figure 39: Family Test - Indonesian Fighters Returning from Syria & Iraq 68](#_Toc507503169)

[Figure 40: Family Test - Indonesian ISIS Fighters Released in Jail Breaks 69](#_Toc507503170)

[Figure 41: Family Test - Indonesia Growth of ISIS 70](#_Toc507503171)

[Figure 42: Family Test - CounterTerrorism Results for Green 71](#_Toc507503172)

[Figure 43: Family Test - Indonesia Baseline CT Results 72](#_Toc507503173)

[Figure 44: Family Test - Indonesia Terrorism Deaths 73](#_Toc507503174)

[Figure 45: Family Test - Indonesian Legitimacy of Green 74](#_Toc507503175)

[Figure 46: Family Test - COA Impact on Green Legitimacy 77](#_Toc507503176)

[Figure 47: Family Test - COA Impact on Green Legitimacy 78](#_Toc507503177)

[Figure 48: Family Test - COA Impact on Total Combatants 79](#_Toc507503178)

[Figure 49: Family Test - COA Impact on Territory Controlled 80](#_Toc507503179)

[Figure 50: Family Test - COA's impact on Detainees Released 81](#_Toc507503180)

[Figure 51: Family Test - Impact of COA's on Terrorist Attacks 82](#_Toc507503181)

**LIST OF TABLES**

[Table 1: Sensitivity of Ethnographic Perception & Transition Parameters 105](#_Toc507503182)

# C-1 Introduction

The following validation tests were run against E-SAM:

* Boundary Adequacy
* Structure Assessment
* Dimensional Consistency
* Parameter Assessment
* Extreme Condition
* Integration Error
* Behavior Reproduction
* Behavior Anomaly
* Family Member
* Surprise Behavior
* Sensitivity Analysis
* System Improvement

# C-2 Boundary Adequacy

For both baseline scenarios, *Historical* and *Without Intervention,* the following boundaries were established:

|  |  |
| --- | --- |
| Boundary Topic | Boundary |
| Geography | Combined geography of Iraq & Syria |
| Ethnographic Groups | Arab Suuni, Arab Shia & Kurdish Suuni populations. |
| State-Sponsored Foreign Intervention [Purple] | Iran, Hezbollah, US backed Coalition, Russia & Turkey on behalf of [Green]. (Only in *Historical.)* |
| State-Sponsored Non-State Actor Interventions [Green] | Kurdish Syrian Defense Force (SDF) & Arab Shia Popular Mobilization Forces (PMF) |
| Time Duration | 10 Years (40 Periods) |
| Strategic Architecture Sectors: | Resource Stocks, AFV/IFV, OpOrder Allocations, Governance, Combatant Recruiting & Losses, Foreign Intervention OpOrder Allocations |
| World Model Sectors: | Sources of Revenue, Sources of Expenses, Territory & Scenario Data, SFS Combat Simulator, Resistance & Uprising, Ethnographic Perceptions, Militant Recruiting & Losses, Actor Legitimacy & Side Choosing, OpOrder Impacts on World, |

## Boundary Tests

***Time:*** The Time boundary was originally selected at 10 years. A boundary test was conducted by expanding the duration to 20 years. Results were compared against the primary measures of *Territory Controlled by Actor[Red]* and *Total Population Controlled by Actor[Red]* in Figures XX and XX.

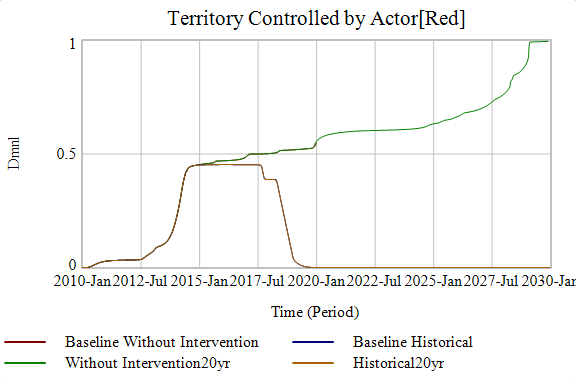


Figure : Time Boundary Test on Territory 20yr

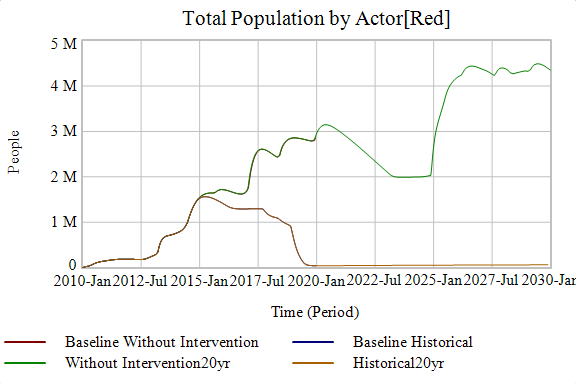


Figure : Time Boundary Test on Population 20yr

Although it appears that the 20 yr model duration in Figure XX allows Red Actor to “win” by conquering all territory, and thus grow its population in Figure XX this is caused by formulation choices and not endogenous system-level dynamics. Although the model has the capability for demographic growth built into the structure, for both baseline scenarios the *Demographic Growth* of all ethnographies is set to 0. This parameter is found in the World Model sector “Actor Legitimacy & Side Choosing.” When this parameter is set to a nominal 2.5% and run for 20 years, what had appeared to be the victory of Red turns into a new equilibrium at similar levels to the 10yr run as seen in Figure XX.

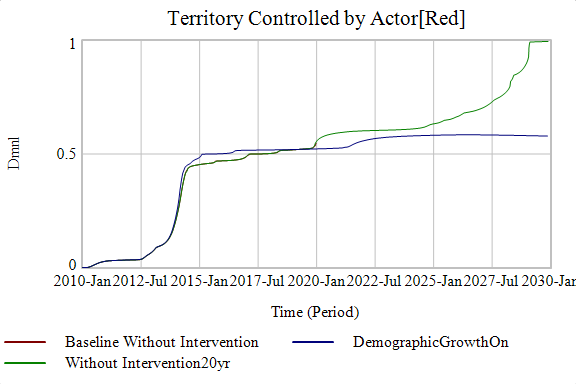


Figure : Time Boundary Test with Demographics Activated

Given these tests a 10year model duration is an acceptable boundary.

***Intervention Tests***

The boundary on what to include in terms of intervention can be examined conceptually by looking at sufficiency. The *Baseline without Intervention* presents the counter-factual “what-if” behavior of Red with no eternal intervention. What is necessary and sufficient to explain the actual historical behavior in terms of intervention? The *Baseline with Historical Intervention* successfully recreated such behavior by including two kinds of intervention responses. A foreign-supported one that included the responses of Hezbollah, Iran, Russia, Turkey and the US etc. And a second local-supported intervention that arose from indigenous populations, including the Kurdish Suuni based Syrian Defense Force (SDF) and the Arab Shia based Popular Mobilization Force (PMF). Two tests were ran – each one excluding one half of this intervention response. The first by removing all foreign interventions named *Local Only* and the second test which removed the local non-state actor interventions but preserves foreign interventions called *External Only*. This second test was called *External Only.* These two tests were then compared to see if only one form of response to ISIS was sufficient, or both required to recreate historically observed behavior. Results were compared against the primary measures of *Territory Controlled by Actor[Red]* and *Total Population Controlled by Actor[Red]* in Figures XX and XX. The secondary measures of *Blue or Purple Intervention Size[Green], Combatants[Arab Shia, Green]* and *Combatants[Kurdish Suuni, Green]* demonstrate the removal of these forces as shown in Figures XX and XX.

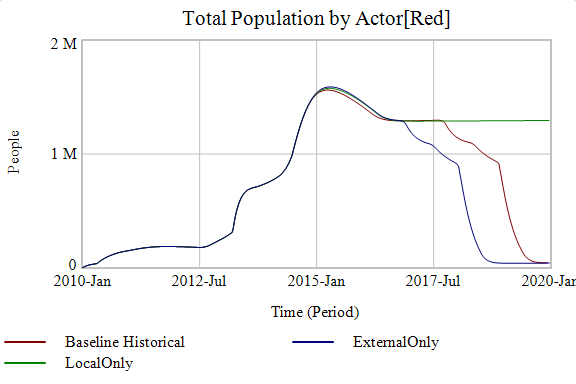


Figure : Interventions Boundary Test on Total Population

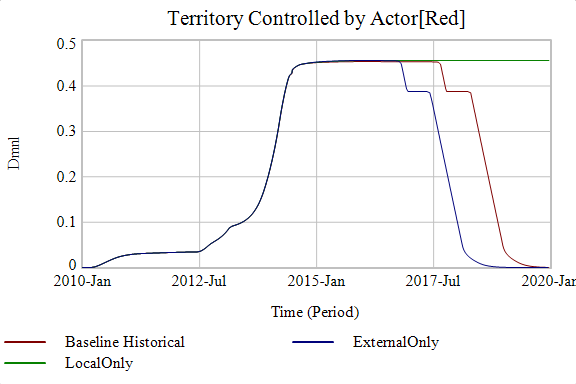


Figure : Intervention Boundary Test on Territory

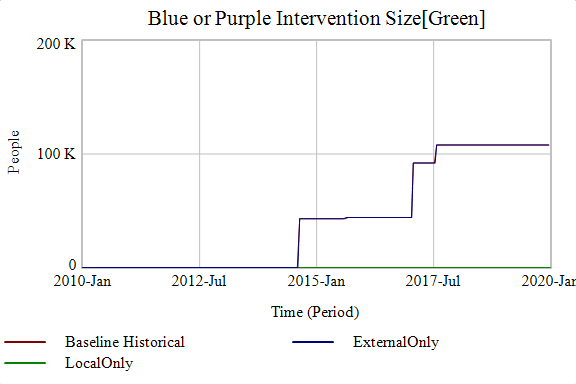


Figure : Intervention Boundary Test External Intervention Sizes

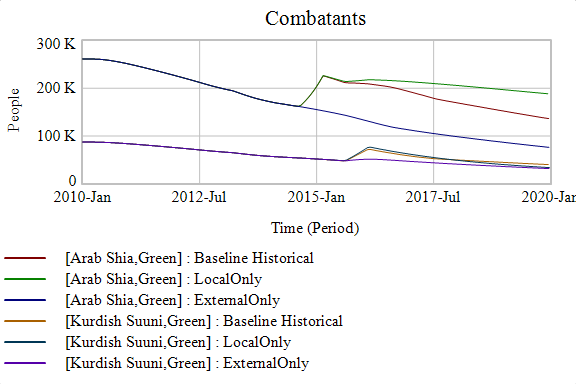


Figure : Intervention Boundary Test - Local Intervention Sizes

The test indicated that *Local Only* interventions were not sufficient to recreate the baseline behavior. The amount of *Territory* the Red Actor controls flatlines and does not decrease as was historically observed. A surprising behavior is that not only does *External Only* recreate the appropriate historical behavior, but it does it sooner than the *Historical Baseline* which combined both types of interventions. This is a counterintuitive result – *External Only* has nearly 100,000 less combatants than the *Historical Baseline.* So why did it perform better?

The cause for this improved performance is found in *Average Combatant Experience[Green]* and how it increases more rapidly in the *ExternalOnly* test as shown in Fig XX.

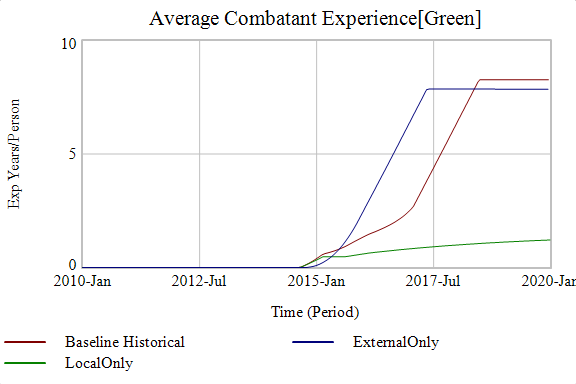


Figure : Intervention Boundary Test - Average Combatant Experience

This is another manifestation of the well understand XX effect. Adding more people to a project slows down progress as new arrivals have to be trained. Foreign soldiers deployed into Syria and Iraq did not require the same training as local civilians who joined militias such as the SDF or PMF. Without the experience drag of these local non-state actors, in the *ExternalOnly* scenario Green combatants are trained faster, have more experience and perform better in combat. This allows Green to take a more aggressive offensive stance sooner in the conflict, resulting in the faster achievement of the defeat of Red as shown in Fig XX.

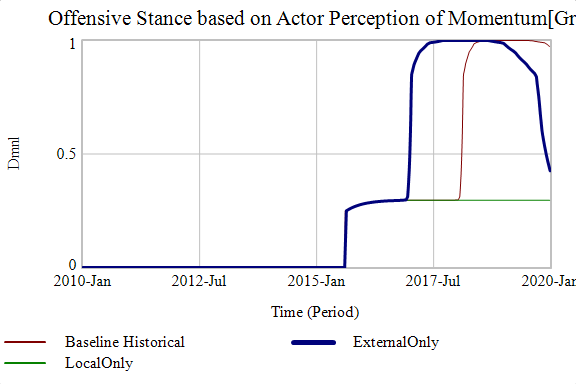


Figure : Intervention Boundary Test - Offensive Stance of Green

Returning to the original boundary test: *ExternalOnly* in isolation has sufficient capability to recreate historical performance while *LocalOnly* does not. However it would be inappropriate to exclude local non-state actor responses when we know historically that they occurred. Therefore the boundary of including both local responses as well as foreign interventions is considered a sufficient boundary.

***Ethnographic Boundary Tests***

The ethnographic boundary for E-SAM in both baselines is selected to include the three dominant ethnographic groups found in both countries: Arab Suuni, Arab Shia and Kurdish Suuni. These ethnographic groups are constructed from combining an ethnic distinction, such as Arab or Kurd, with a religious denomination affiliation such as Suuni or Shia. All other ethnographic groupings eincluding the ethnicities of Turkomen and Assyrian or the religious affiliations of Druze, Yazidi, Christian etc. are excluded by this boundary selection. The reasoning is that within the selected ethnicities 90-95% of the population is represented, and a similar coverage is obtained within religion.

The E-SAM model can handle any number of ethnographic groups because they are subscripted. However, because the research questions answered by the *Historical Baseline* and *Baseline without Intervention* are not focused specifically on ethnographic performance under conflict and is rather focusing on the theater level conflict it is not plausible that these small minorities would have significant impact on the conflict that wasn’t already represented by the behavior of one of the three main groupings. For example Turkomen Suuni might be both targeted for recruitment by Red even as their ethnographic group is persecuted, but at a lower level than Arab Shia. These circumstances are already captured in the Kurdish Suuni ethnographic group.

Therefore the ethnographic boundary of only including three groups is considered plausible since adding additional groups would not significantly alter the outcome. For research questions specifically targeting ethnographic performance in periods of conflict, such a refugee status, additional ethnographic distinctions can be added as needed.

# C-3: Structure Assessment

As discussed in Precision vs. Realism section in the overview, structural assessment was a primary means of calibration over numerical payoff optimization.

### Conservation of Mass Errors

Earlier versions of the model at times resulted in negative ethnographic populations under various circumstances. This was a function of having multiply independently calculated outflows to the same stock. Some of these were ratio outflows, for example Governed Population transitioning to Calculated Legitimacy. Others were integer outflows produced by other sectors in the model: civilian deaths due to terrorism or war crimes, refugees fleeing the area. Though each outflow itself had first order negative control – there was no overall first order control that could govern all the outflows at the same time. Figure XX demonstrates an example of this problem when *Historical Baseline* is compared *Historical Population Conservation Error* on the primary measure of *Total Population[Red].*

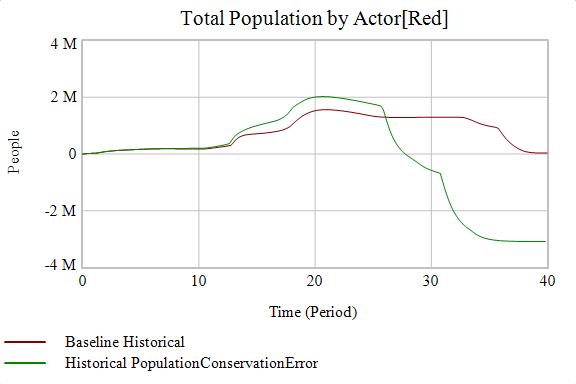


Figure : Structure Assessment - Conservation of Mass Error

This class of structural errors was solved by creating a single first-order control called *Ethno by Actor Sufficiency* displayed in Figure XX.



Figure : Structure Assessment - Conservation of Mass Correction

The formulation of this structure is to compare the current ethnographic population versus a reference number, and apply a lookup function that reduces the sufficiency by the resulting ratio. Conceptually as a population of an actor becomes scarce – the effectiveness of any act against them reduces as they become harder to locate or target. As the actual population dipped beneath the reference level a reduction percentage is calculated and applied to the outflows preventing populations from going below zero due to civilian deaths, refugees and civilians lost to conquest. The values of these modifiers for *Historical Baseline* are shown in Figure XX.

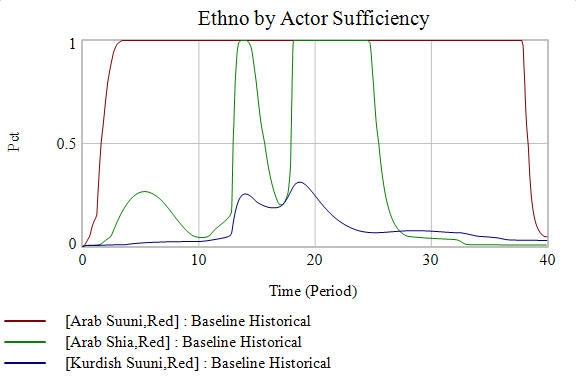


Figure : Structure Assessment - Ethno by Actor Sufficiency Values

### Free Lunch Errors

Another class of errors are called “free lunch” errors. An actor continues to undertake an activity even if they don’t have the resources or means to plausibly do so. This is demonstrated by comparing the *Historical Baseline* to *Historical Free Lunch Error* across the primary measures of Territory, Total Combatants and Terrorist Attacks in Figures XX, XX, and XX.

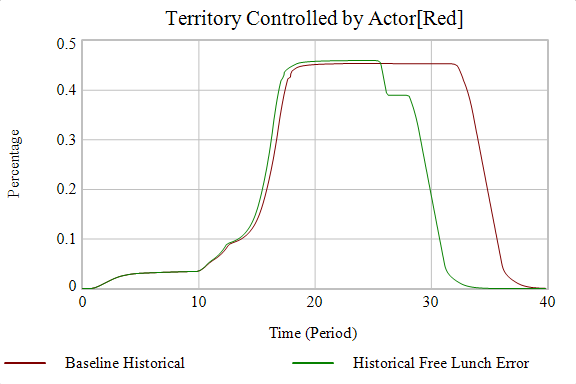


Figure : Structure Assessment - Free Lunch Error Territory

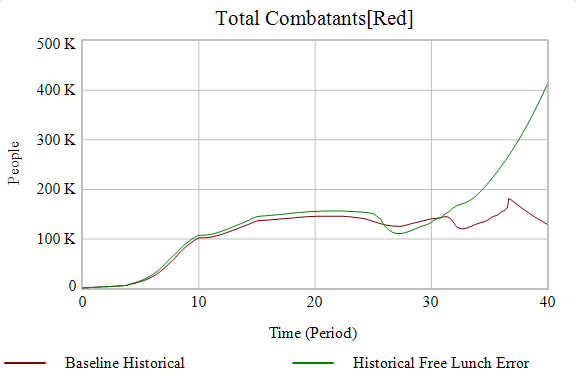


Figure : Structure Assessment - Free Lunch Error Total Combatants

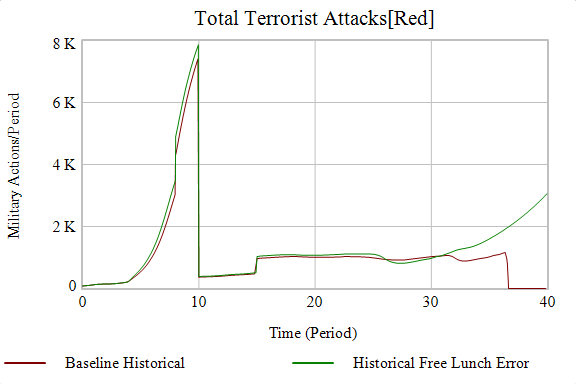


Figure : Structure Assessment - Free Lunch Error Total Terrorist Attacks

It did not seem plausible that Red actor could lose all its territory and yet experience significant increases in *Total Combatants* and *Total Terrorist Attacks.* This wouldn’t match the historically observed behavior. Although ISIS certainly continues to conduct terrorist attacks and recruit combatants outside of Syria and Iraq, E-SAM is bounded to Iraq and Syria only precluding those as logical explanations.

The error was found in the financial sector of the Strategic Architecture. In previous versions monies accumulated above and beyond what was necessary to conduct operations were stored as reserves. This excess was described as likely being sent to actors abroad, but since the model was focused on Syria and Iraq, structure was not added to explicitly demonstrate this spending abroad. As a result even when all Territory, and with it population derived resources and territorial derived resources eliminated, Red Actor still had significant reserves from which to operate from as shown below in Figure XX.

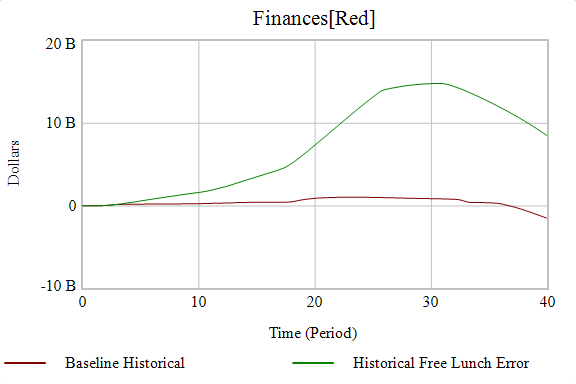


Figure : Structure Assessment - Source of Free Lunch Error

If funds are not sent abroad, Red Actor has a significant surplus of reserves built up it can continue to use even after it has lost all Territory. The solution was to add structure realistically replicating Red Actor sending money overseas to fund different efforts. However another problem quickly arose that the model had no means for either Actor to prioritise the funding of essential or non-essential tasks. This was considered a sufficient level of aggregation in earlier versions of the working draft. But as Red Actor ran out of funds, it became apparent that they would realistically begin to cut non-essential payments: such as detention benefits to ISIS members in prison and death benefits. Ultimately essential services would have to be cut as well as bankruptcy loomed. Additional structure – both to allow Red to send money abroad and to prioritize between essential and non-essential payments was added as shown in Figure XX below.



Figure : Fixing Free Lunch Errors - Sending Funds Abroad & Spending Prioritization

This simple structure allowed for ISIS to spend money abroad when times were plentiful, restrict sending money abroad as funds dried up, then begin to cut first non-essential payments and ultimately essential payments as they headed towards bankruptcy. When Red Actor doesn’t have funds it can’t pay for military actions and can’t continue to recruit or conduct terrorist attacks, creating more realistic – and historically accurate behavior.

A cascading benefit of this new structure is it enabled a fix of two other structural assessment problems. Although *Total Defections* were already calculated using the ethnographic perception of the actor, this wasn’t sufficiently depleting *Combatants* through defections as was historically perceived. Likewise, Red Actor *Detainees in Prison*, simply accumulated over time to massive levels allowing unrealistically high inflows of *Detainees Released* through *Prison Breaks* to rejoin the Red Actor force. These unrealistic By adding structure to reduce payments in response to budget pressures to non-essential (*Detention Benefits)* and essential budget reductions to essential (*Payroll)*  the gap between desired and actual payments in these areas could be considered in additional *Defections* from either the active, or imprisoned, ranks of an Actor. Figure XX shows the structure added in the Expenses sector of the Strategic Architecture, Figure XX structure added to *Total Defections*  from *Combatants* and Figure XX structure added to *Detainees in Prison.*



Figure : Fixing Free Lunch Errors - Allocation of Essential & Non-Essential Budget Priorites to Payroll and Detainee Benefits



Figure : Fixing Free Lunch Errors - Payroll Gap Structure added to Total Defections Rate



Figure : Fixing Free Lunch Errors - Detention Benefits Structure added to Defections within Prison

These three additional structures generated the following behavioral changes in the model. First gaps emerged between desired and actually paid *Payroll*  as well as *Detention Benefits.*

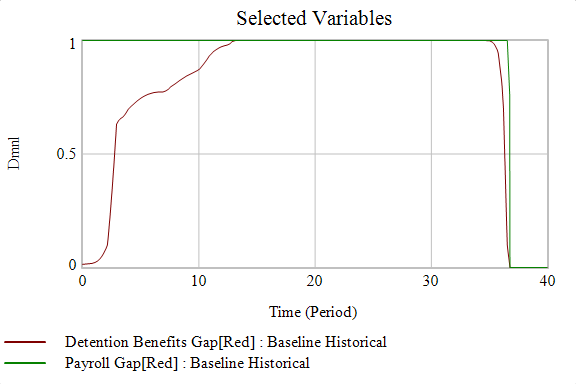


Figure : Fixing Free Lunch Errors - Gaps based on prioritization of Spending

As this information reached *Combatants* and *Detainees in Prison* respectively, the defection rates for both adjusted accordingly. Especially when compared to behavior without these structural changes as shown in Figure XX for *Total Defections*.

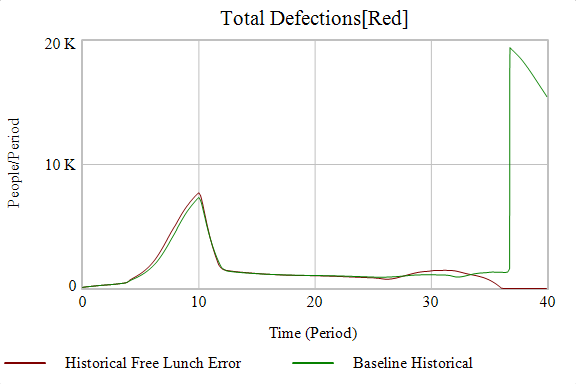


Figure : Fixing Free Lunch Errors - Comparison of Total Defections Rates

Although there are many reasons a combatant might abandon ISIS other than a gap in payroll including ideological, tribal, psychological pressures – using payroll gap was considered sufficient to aggregate this behavior. As ISIS begins losing to the point it can no longer pay it’s troops – they begin defecting. Likewise where prior to this structural change there was no way to reduce the *Detainees in Prison* except by prison breaks, with the change imprisoned ISIS followed naturally defect endogenously as their benefits stop getting paid.



Figure : Fixing Free Lunch Errors - Comparison of Detainees in Prison

These structural changes resulted in a much more realistic behavior of the model. When ISIS lost all its territory, it lost it’s ability to generate resources, this led to first a reduction in sending money abroad, then in spending non-essential payments and finally suspension of essential payments. As a result they could no longer afford to perform *Military Actions*  even as defection rates from active *Combatants* and *Detainees in Prison* increased. All based on endogenous feedback rather than parameter calibration.

### Conservation of Information Errors

A final class of errors discovered via structural assessment tests was the conservation of information. In the original E-SAM models there was very little in the way of “leadership intelligence” on how an Actor viewed themselves in relative competition to the opposing Actor. If Red Actor had money, they would fight to expand their territory. Green Actor would defend. But there was no point Green Actor was able to “realize” that the expansion of Red was beginning to slow and thus adopt a more offensive posture of their own forces. This is demonstrated in Figure XX on three different versions of the model where the *Baseline Historical* is compared against an *Error Type 1* and an *Error Type 2.*

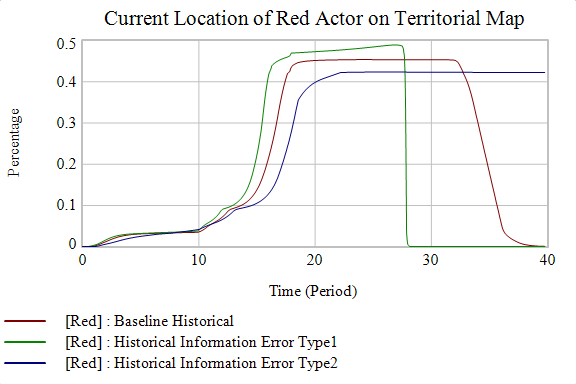


Figure : Structure Assessment - Information Error

In all three scenarios the same foreign and local interventions occur. In Error Type 1, the shift in momentum is binary – resulting in a historically unrealistic immediate defeat of ISIS. In Error Type 2, even though Green has brought Red forward progress to a halt, Green doesn’t “realize” this and shift to a more offensive posture. Resulting in an unrealistic stalemate.

This structural error was corrected by adding ‘intelligence’ at the leadership level of each Actor as shown in Figure XX.



Figure : Structure Assessment - Correction to Information Error in Perception of Momentum

The structure takes as an input changes in territorial control and tracks as a stock *Perception of Territorial Progress.*  However, an average of the rate-of-change in that stock is then an input to another inflow to a stock which tracks *Perception of Momentum.*  This stock then converts into an *Actor Perception of Momentum* which is the leadership understanding if they are winning or losing, territorially, and by how much. This perception is then compared to a lookup function which converts perception into an *Offensive Stance.* The more an Actor perceives themselves to be winning – the higher their *Offensive Stance.*  The concept of *Strategic Surprise* also plays an influence. An actor subject to strategic surprise, which can be determined by scenario – is not going to suddenly shift offensively as they are still “remembering” the impact of surprise. The result of this structure is a “posture” that the Actor takes in terms of allocating offensive actions based on their perception of how the conflict is going, subject to surprise. This allocation is further modified by how much territory the Actor has remaining – they are willing to commit more forces to defense as they lose more territory than when the opponent is taking away farther portions from their center. The *Allocations* are shown below for both Green and Red Actors for the Historical Baseline (Figure XX).

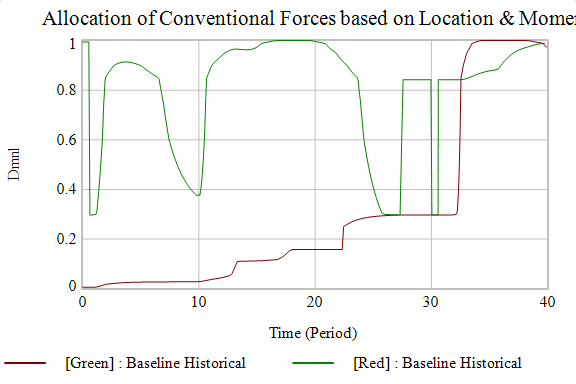


Figure : Structure Assessment - Allocation of Conventional Forces Historical

Importantly, the new structure results in independent assessments of momentum by each Actor. As Red Actor struggles to progress during drawn out city-sieges or hostile ethnographic terrains – they pull back a little on their offensive allocation. Meanwhile, as Green Actor perceives the progress of Red Actor diminishing they begin gradually increasing their willingness to go on the offense until final, just after period a tipping point is reached and Green goes on a full offensive. The reason Red Actor remains at a high allocation is because of the aforementioned desire to increasingly defend the last pieces of territory they possess. This is contrasted with the Baseline Without Intervention scenario as depicted in Figure XX.

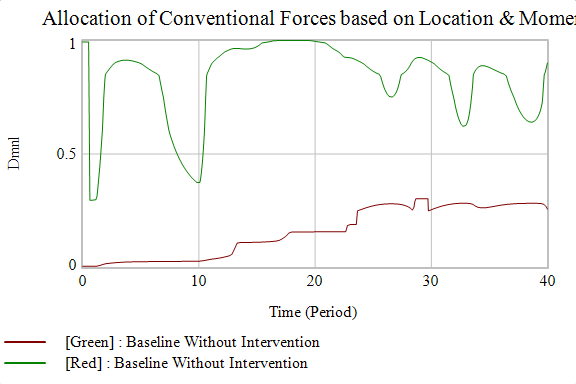


Figure : Structure Assessment - Allocation of Conventional Forces without Intervention

In this scenario although Red experiences varying perceptions of their own progress, Green never accumulates a confidence that it is beginning to win. Although they do increase their offensive allocation, it is at a lower rate. Note that these perceptions of winning and losing endogenously create a natural equilibrium when the two sides reach a stalemate point.

This kind of structural assessment to correct an information error can also be found in how Ethnographic populations in the Unaligned category pick between two sides on who to join. This structure is depicted below in Figure XX.



Figure : Structure Assessment - Correction of Information Error in Ethnographic Perception of Momentum

The information error in this case was the ability for an ethnographic group to perceive the relative momentum between a State and Non-State Actor in terms of Calculated Legitimacy, represented by *Pct views Actor as Best Choice for Now.* Each separate Ethnographic group evaluates this relative momentum to determine – not who is the best for them – but who appears to be getting better and who appears to be getting worse in terms of legitimacy. This then fuels the rate of change out of the *Unaligned Population* stock and into the *Calculated Legitimacy* population stock for either Actor. It represents a realistic side-selection of those willing to take a gamble that current conditions indicate a better future by aligning themselves with a certain actor. The results of this information error correction structure are shown in the chart below which depicts *Ethnographic Relative Momentum Percpetion* in Figure XX for the *Historical Baseline.* Note that negative perceptions are in favor of Unaligned converting to Red, and positive perceptions are in favor of Unaligned converting to Green.

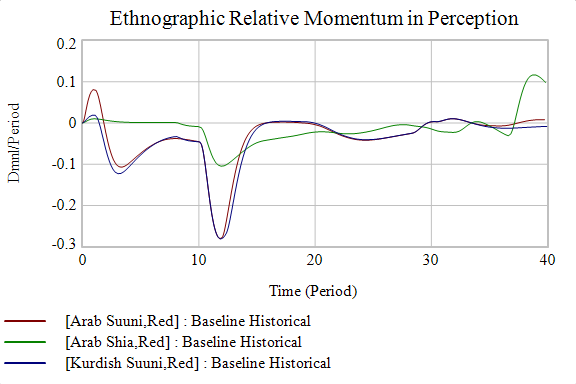


Figure : Structure Assessment - Relative Ethnographic Perception of Momentum Historical

It’s important to remember that this is a perception. The rational observer in retrospect might find no circumstance under which it was plausible for *Kurdish Suuni* or *Arab Shia* to ever desire to join ISIS. However, the model is treating ethnographic groups as agents with bounded rationality. They only know what they can perceive, weighted to recent events. Arab Shia do indeed show less willingness to convert to the Red Actor due to their persecution, and more willingness to convert back when the tide seems to be shifting. Compared to the *Baseline without Intevention* below where the two sides reach a stalemate of perceived momentum in serving any ethnographic group as shown in Figure XX.

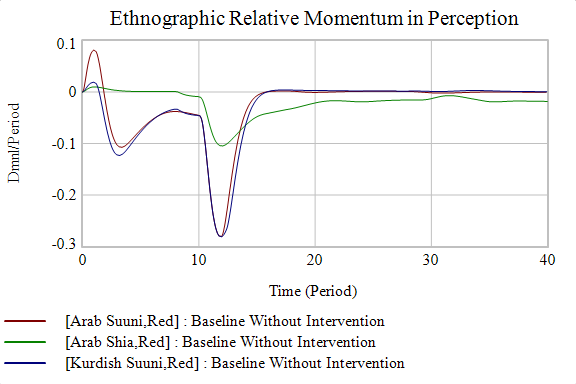
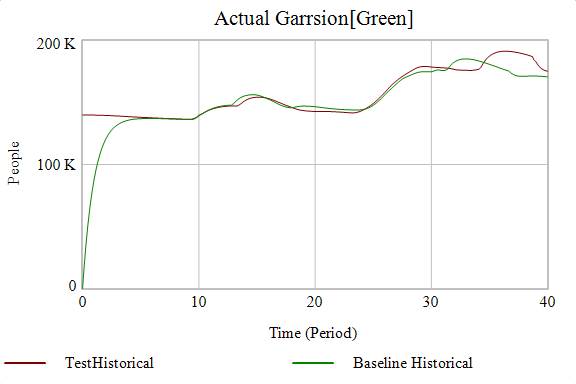


Figure : Structure Assessment - Ethnographic Perception of Relative Momentum without Intervention

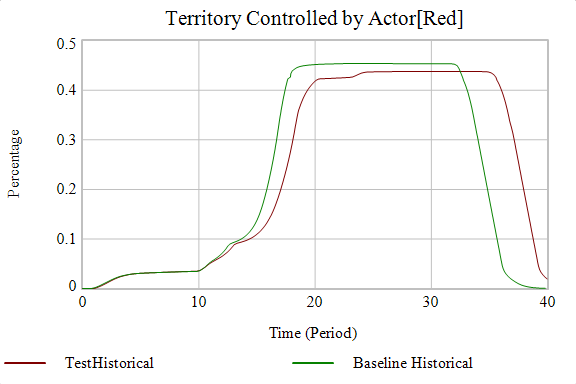
### Formulation Reviews

In some cases even if a structural concept was plausible and sound – the formulas within the structures were flawed creating unrealistic behaviors. Two examples of this were the calculation of *Actual Starting Garrison* and the *Initial Ethnographic Perception.*

In early testing, *Actual Garrison*  was initialized through a computation of existing model structure. This created a simultaneous equation error and to avoid was set at zero. This created unrealistic behavior where the *Actual Garrison*  would have to “adjust up” to a normal level which had consequences on the rest of the model. This was fixed by manually taking the initial *Desired Garrison & Police Forces* and manually setting that as a separate *Actor Starting Conditions Initial Garrison* removing the simultaneous equation problem and creating more realistic behavior. This is shown below in Figure XX, which compares *Actual Starting Garrison* in a Baseline Historical (before fix) and Test Historical after the fix.

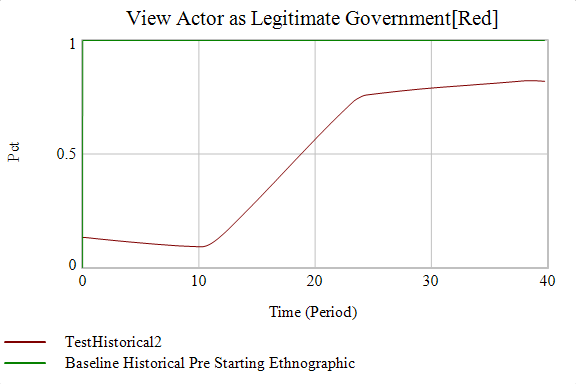


It is implausible that there were no troops conducting garrison or law enforcement so the Test Historical with formulation fix is more realistic. This change impacts the Primary Measure of Effect *Territory Actor Controls* as shown in Figure XX.

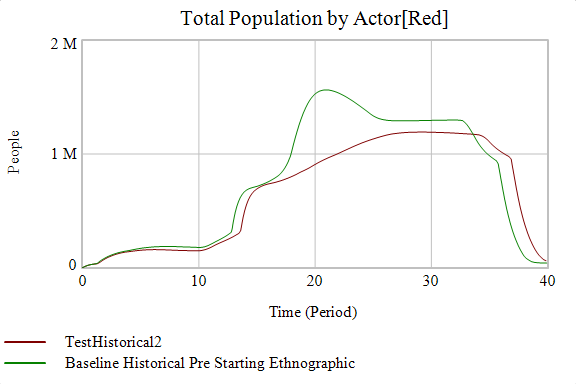


Without having to “ramp up” a Garrison in the Test Historical Red performs slightly worse with a later breakout and corresponding decline.

A second example of using formulation review to structurally calibrate the model was in how Initial Ethnographic Perceptions, both current and the generational anchor, were calculated. As a result in earlier runs of the model Red was starting off and reaching 100% Legitimacy almost immediately, which is not realistic. This is demonstrated below in Figure XX which compares *View Actor as Legitimate [Red]* pre fix and post fix.



The impact of this formulation error was that Red was able to convert population into Governed, gaining more benefits, faster than was realistic. When the formulation is fixed a slightly more realistic behavior appears in *Total Population by Actor [Red]* indicating as shown in Figure XX.



This is just a sample of all the activities performed under structural calibration. But they demonstrate how calibration by structural assessment and formulation review helped improve the “fit” of the model to historical behavior without numerical optimization of parameters to achieve a higher payoff via brute computation. Structural calibration also operates in environments where numerical “fit” cannot. For example, the actual population controlled by ISIS historically may never be known. This would prohibit even the attempt to numerically calibrate a model. But structural calibration continued to improve the ‘fit’ of behavior patterns relating to population by eliminating implausible structures and errors in formulation.

In the following Figures the effects of this structural calibration effort are demonstrated in several primary measures of effect. Note that in all examples “Test Historical 2” is the most realistic ‘fit’. Figure XX demonstrates the different results throughout the structure assessment on the Primary Measure of Effect on *Territory Controlled by Actor.*

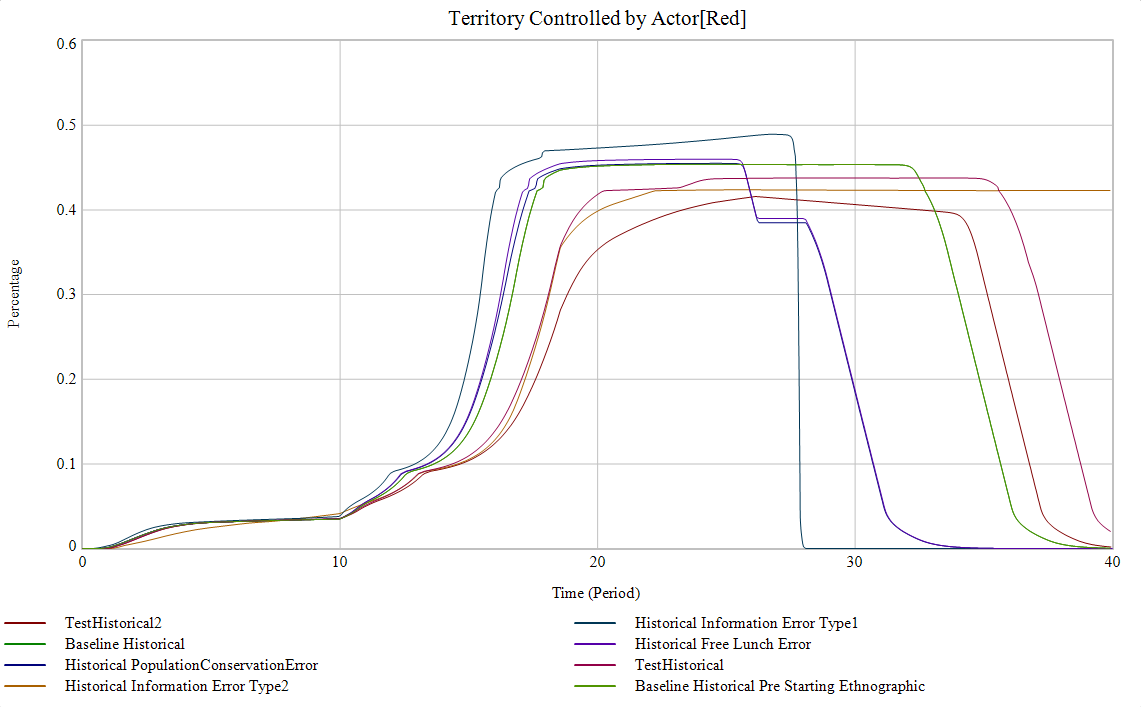
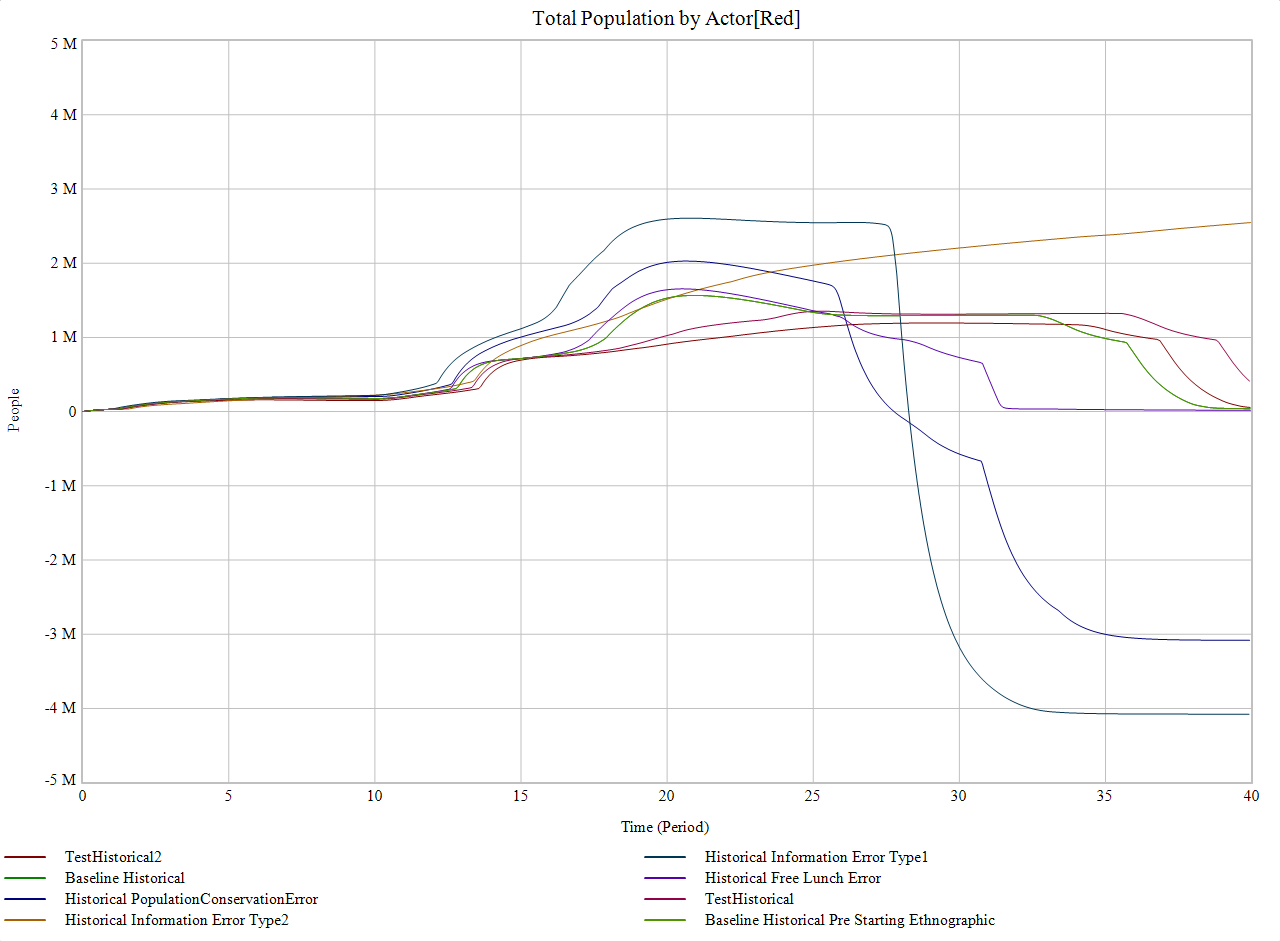
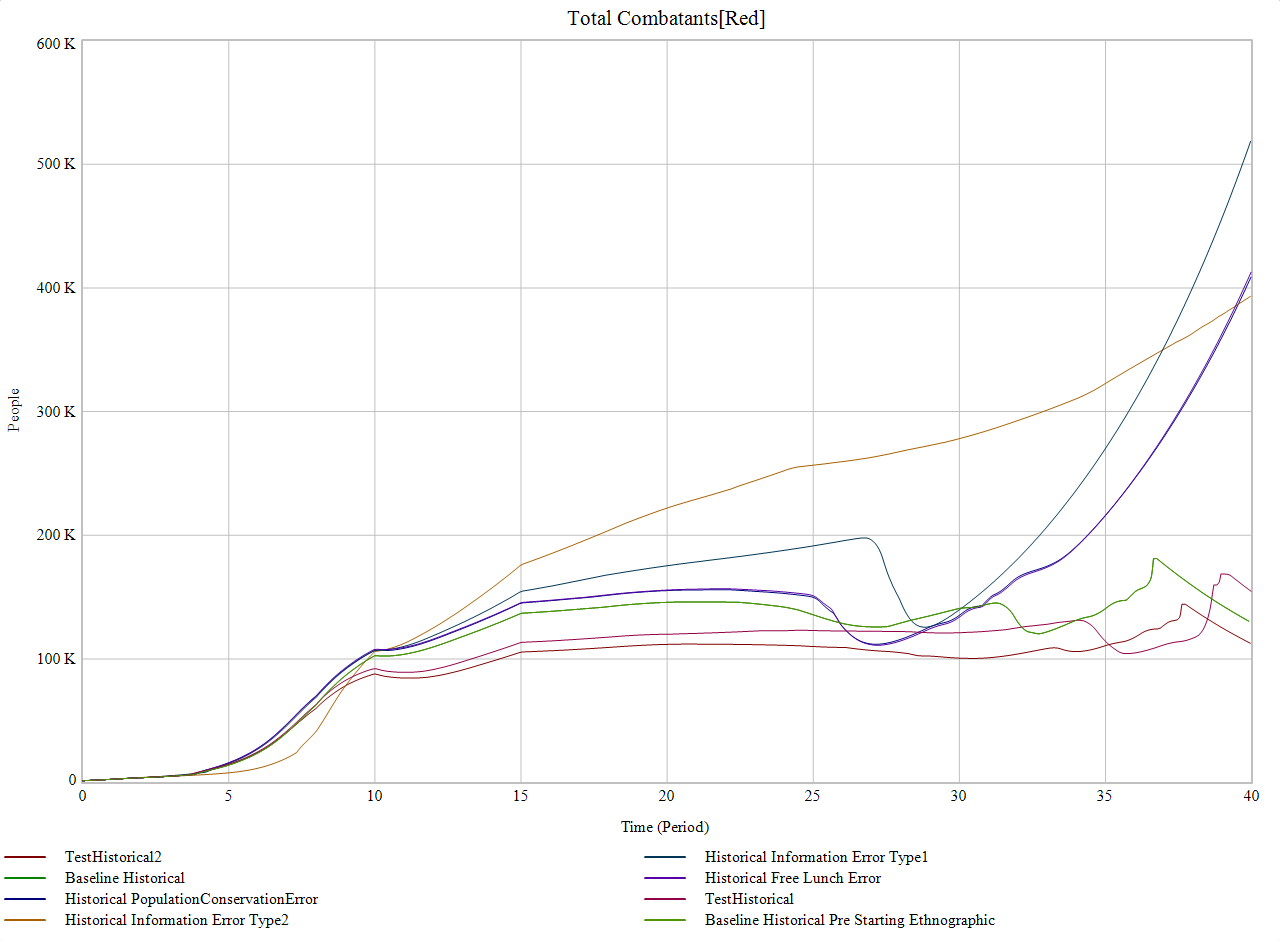


Figure XX demonstrates the same on *Total Population by Actor.*



And Figure XX on *Total Combatants.*



STRUCTURE CALIBRATION CHART

Test Historical 2

Test Historical

Baseline Historical Pre Starting Ethnographic

Baseline Historical

Historical Information Error Type 1

Historical Information Error Type 2

Historical Free Lunch Error

Historical ConservationError

In conclusion since figures like the actual population controlled by ISIS may never be known – it doesn’t mean that structural calibration resulted in more “accurate” parameter estimates that drive Red’s population in the model. But by focusing on structural assessment and formulation review errors were uncovered that created clearly implausible or unrealistic behavior in subsystems of the model. Had numerical calibration been attempted instead, these errors might have been overlooked in favor of parameter adjustments to find the right “fit.” This would virtually ensure that the parameters which displayed the best fit were clearly wrong as they included what are now known to be structural and formulation errors. Although we may never know if the parameters found in the model are accurate – through structural calibration confidence has been increased that the structure and formulations within which those parameters reside are realistic building confidence in the model results.

# C-4: Dimensional Consistency

The dimensional consistency of E-SAM is checked both by software and manually inspected. The software review returns a value of acceptable units consistency as shown in Figure XX.

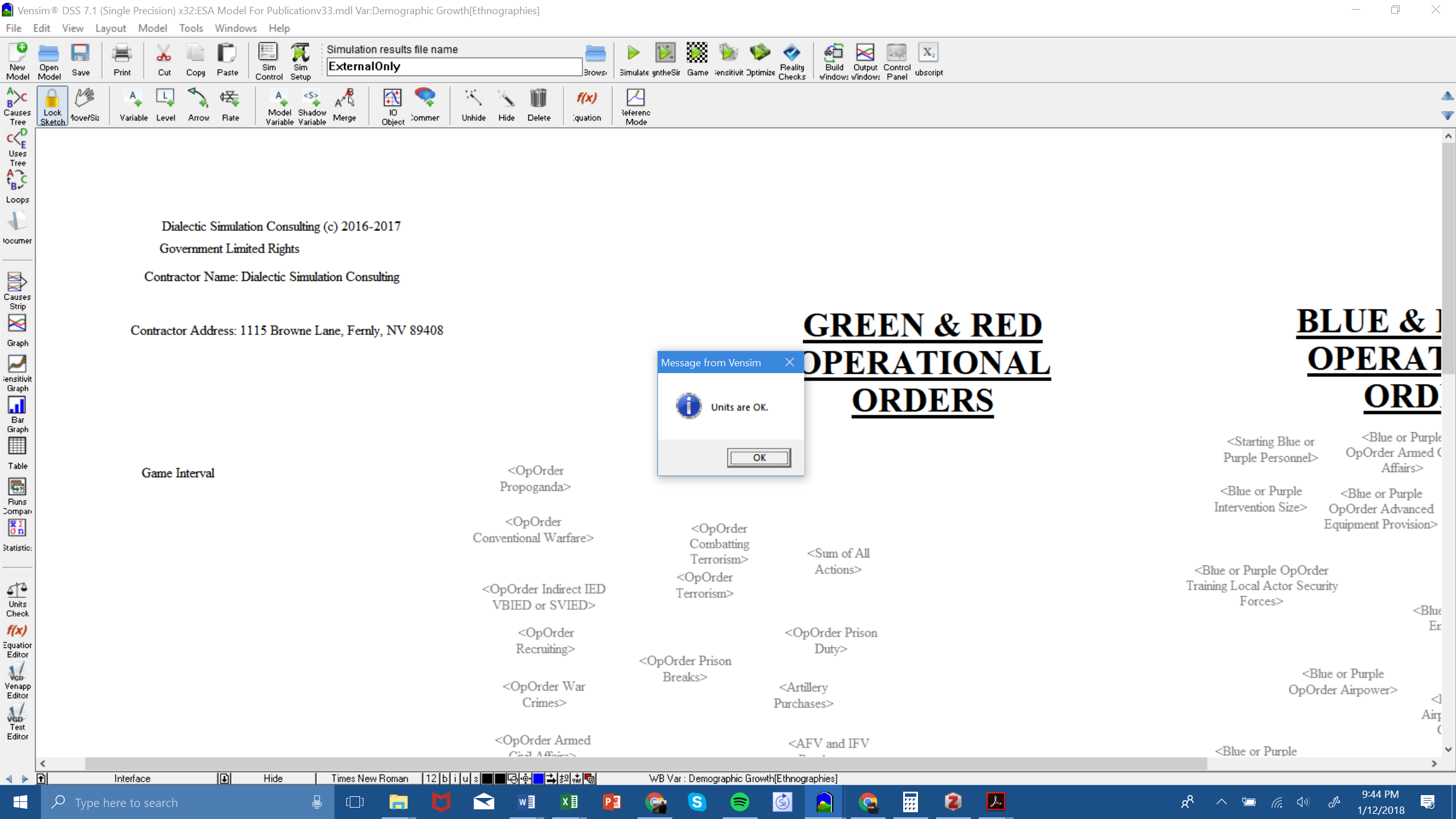


Figure : Unit Consistency

# C-5: Parameter Assessment

The source of parameter estimations is covered in the sector-by-sector overview. As noted in the *Precision vs. Realism* section in the Overview, the focus was made on obtaining real-world data where applicable. Unfortunately as ISIS was a fast moving phenomena and much about them is either unknown or classified most of these parameter information came from publicly available, unclassified or declassified sources. Where numerical estimates were not available – modeler judgement was used to create causally realistic parameter values.

However E-SAM itself provides a framework for future parameter assessment. As research is undertaken and published that isolates and focuses on the specific parameters which build the model – these parameters can be updated for E-SAM. A historical implementation of E-SAM may have more robust parameterization benefitting from the distance of time – though historical conflict also won’t have the kind of numerically robust research associated with it outside of the field of cliodynamics that modeling often asks for. Time delays specifically, which have a large impact on the speed at which behavior develops and emerges – may be very difficult to ever formally quantify outside prudent modeler judgement and use of correct structures.

As future versions of E-SAM are updated newly available verified parametric data will be incorporated.

# C-6: Extreme Condition

The E-SAM performs robustly under extreme conditions. Many errors of conservation were identified in structural assessment and while creating the model. However simulation tests at a system level can be performed by creating unrealistic extreme conditions and ensuring the model behaves realistically. For example, a billion combatants joining either side should bring the conflict to a rapid, but not immediate conclusion. Likewise, the mere introduction of such a large force should have second-order consequences in addition to providing a temporary combat advantage – how to pay them all for instance. To test extreme conditions behavior two unrealistic extreme conditions will be introduced. First a billion *Combatants* join Green and Red respectively in the *Baseline Without Intervention* scenario at the start of 2015. This is done by means of a gaming test input *Test Extreme Conditions* which allows a direct addition to the inflow of incoming combatants as shown below in Figure XX.



Figure : Extreme Condition - Test Structure

Two lines are added to the simulation game script:

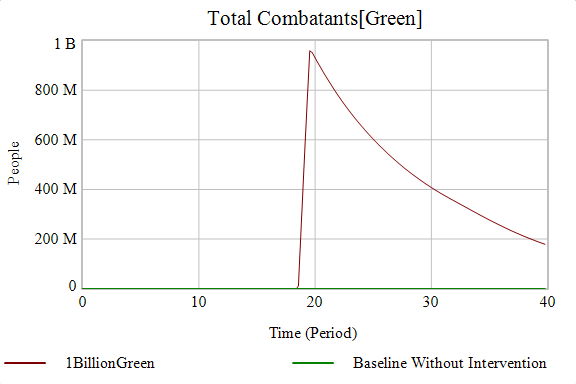
:Time=18.557

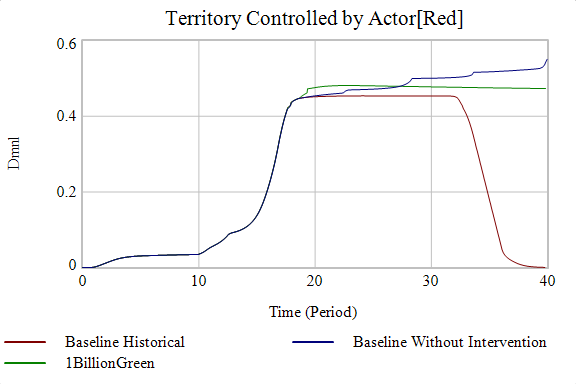
Test Extreme Conditions[Green]=1000000000

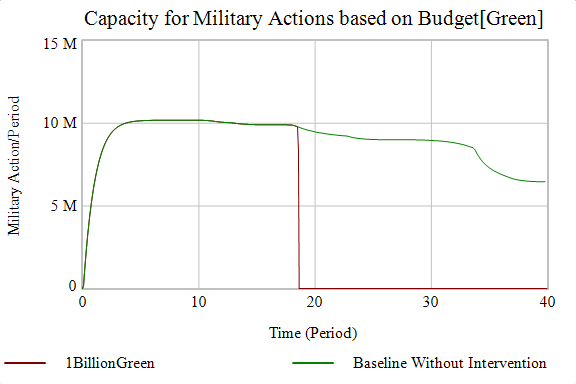
:Time=19.5539

Test Extreme Conditions[Arab Shia,Green]=0

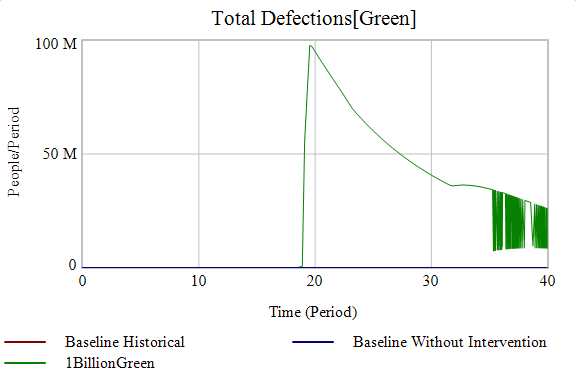
When executing the script we see first that the one billion combatants do join the total force for Green in Figure XX. But when examining the primary measures of effect for *Territory Controlled by Actor [Red]* we do not see expected results of the fight being quickly concluded in Figure XX.



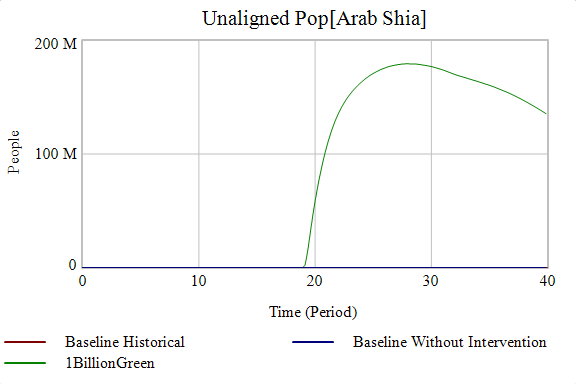


It appears that despite adding a billion extra fighters, Green reaches a stalemate with Red, rather than defeating it. Yet we know from the *Baseline Historical* that Red can be defeated with the intervention of only 150,000 more combatants. Although this at first appears counter intuitive the model is producing realistic results that ripple throughout the sectors. For example, although a billion fighters were added – there was no additional budget provided to pay for this force. This results in a sharp reduction in Green’s ability to conduct offensive military actions as Green goes bankrupt demonstated in Figure XX. 

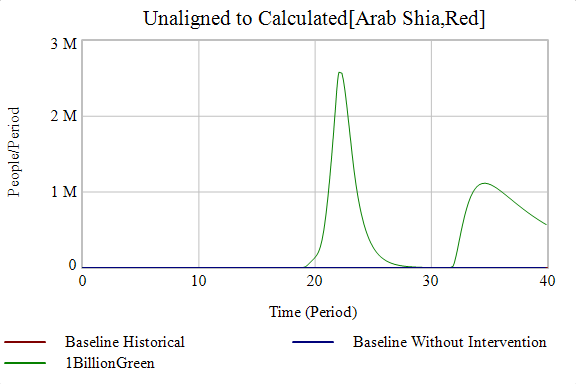
As described in the sector-by-sector overview Green, does not need to use Military Actions to defend Territory with Garrison- which still occurs. But they cannot conduct offensive actions without funds to pay for them. This creates territorial stalemate. But the extreme conditions actually demonstrate robustness in a number of additional ways which are described below as the ripple effects of adding a billion troops to a Territory begins to take hold. First is that Green cannot afford to pay the wages of its military force, resulting in increased defections as shown in Figure XX.



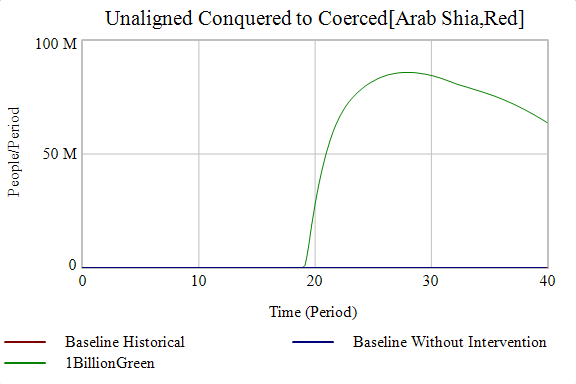
In the model defecting *Combatants* don’t just disappear, they return to civilian life as *Unaligned Population* which will then make a choice of which side to join. The large increase in Unaligned Population resulting from mass defections of Arab Shia from Green is shown in Figure XX.



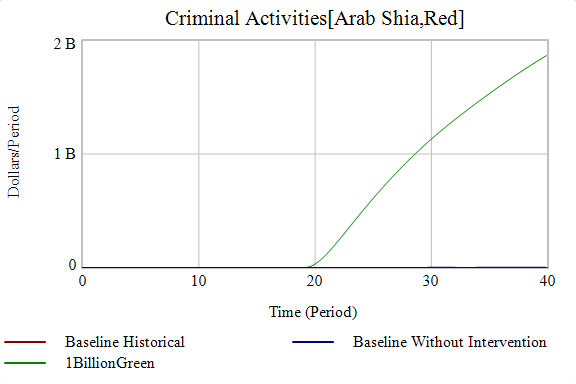
Because at the time the billion combatants are added the perception of relative momentum is working against Green, some of these former *Combatants* but now *Unaligned Population* begin choosing the Red side as shown in Figure XX.



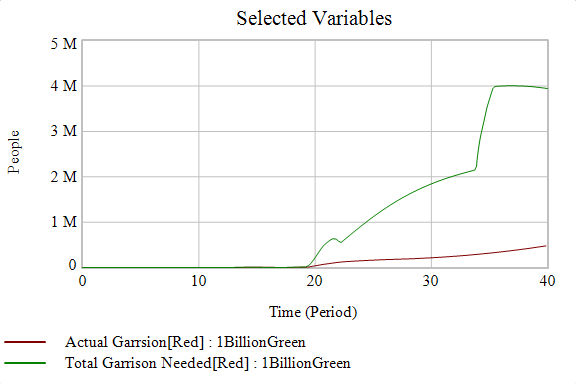
In the model Unaligned are distributed linearly throughout Territory. If Red conquers 50% of the Territory, it will have conquered 50% of the Unaligned. Although this isn’t precise it does reflect that unaligned populations may be mobile – and Red indeed captures a large amount of unaligned Arab Shia which it adds to its own *Coerced Population* in Figure XX.



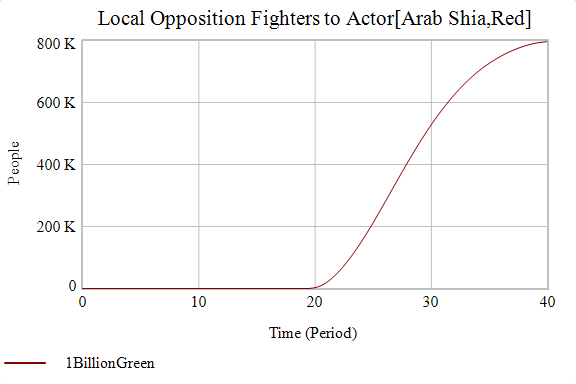
This unrealistic influx of population caused 2nd and 3rd order effects for Red. Although *Coerced* populations do not pay taxes, they can be criminally exploited and this allows ISIS to collect massive revenue from the new Arab Shia population.



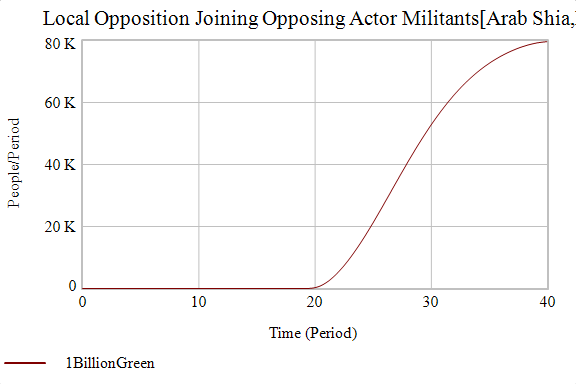
This actually provides the funds for ISIS to continue operating while Green, which brought the people in as *Combatants* was bankrupted immediately. However, because of Arab Shia distaste for ISIS, and that this entire population is *Coerced* Red rapidly runs into problems of providing sufficient *Actual Garrison.* Figure XX shows the large increase in *Total Garrison Needed* versus the *Actual Garrison* that Red can provide based on its available manpower.



Because of its inability to sufficiently garrison this population, Red quickly faces an increase in *Local Opposition Fighters to Actor*  from among the Arab Shia population as shown in Figure XX.



This creates the second component of stalemate – ISIS is too busy fighting an uprising twice the size of the normal Green *Total Combatants.* And from these *Local Opposition Fighters,* some combatants will end up rejoining Green as *Combatants* again as shown in Figure XX.



In summary the extreme conditions test is unrealistic by premise – but creates robust and realistic reactions as it causes ripple effects throughout the model. A billion Arab Shia *Combatants* cannot be paid for by Green, so there is little to no offensive benefit. Unpaid *Combatants* begin defecting to *Unaligned Population* as civilians leaving the battlefield – some of whom are attracted to Red but many more of which are conquered by Red. This causes problems for Red for although it can exploit criminal revenue from this massive mobile population, it does not have the means to provide sufficient *Actual Garrison.* As a result *Local Opposition Fighters* to Red increases in an outbreak of conflict and strife. The inability for Green to pay for additional offensive military actions, and Red’s need to spend all of its *Combatants* in a failed effort to prevent internal rebellion results in the territorial stalemate.

### Billion Combatants & TenTrillion Dollars

A second extreme test is to ask what would happen if the billion Arab Shia *Combatants* were matched with sufficient funds to pay them such that they would not defect and military operations could be conducted? This can be tested by adding *Revenue* of ten trillion dollars at the same time the one billion *Combatants* are added via a *Test Extreme Conditions Revenue.*

**

In the game script the following parameters are added:

:Time=18.557

Test Extreme Conditions Combatants[Arab Shia,Green]=1000000000

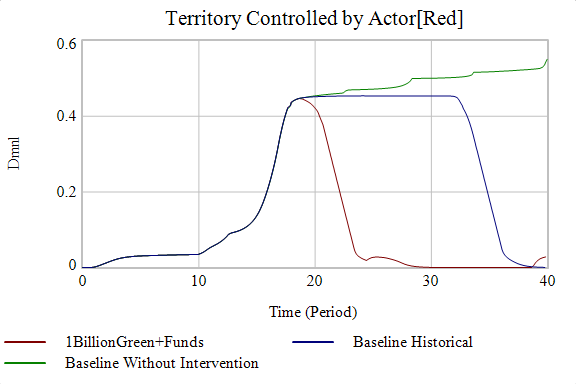
Test Extreme Conditions Revenue[Green]=10000000000000

:Time=19.5539

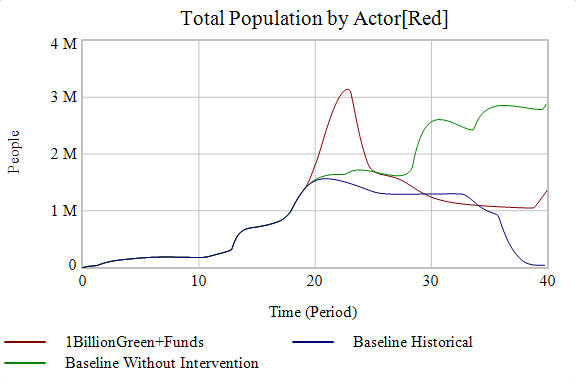
Test Extreme Conditions Combatants[Arab Shia,Green]=0

Test Extreme Conditions Revenue[Green]=0

When sufficient funds are added to the extreme tests the results are what we might predict would happen if one billion equipped and funded *Combatants* entered a conflict – the conflict would quickly end. This is demonstrated by looking at the primary measure of effect *Territory Controlled by Actor* in Figure XX below and *Total Population by Actor* in Figure XX.



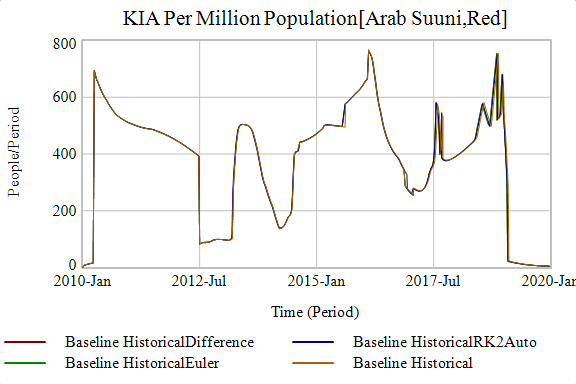
Now that the billion combatants can be paid, and there are sufficient funds to conduct military operations, the Green actor is rapidly able to decrease Red’s territory to zero. Note that it doesn’t happen over night, and still takes time. This is realistic that one billion combatants can’t be everywhere at once – battles have to take time to finish and troops to recover etc. But Red does lose territory. Outside of the military success however some of the ripple effects of introducing a billion *Combatants* to the theater persist even when funds are available. Again this is a demonstration of robustness – defections will still happen at some rate, those defecting will become civilian population and able to pick sides etc. This is why in Figure XX below there is still some population for Red – even though the military conflict has ended.



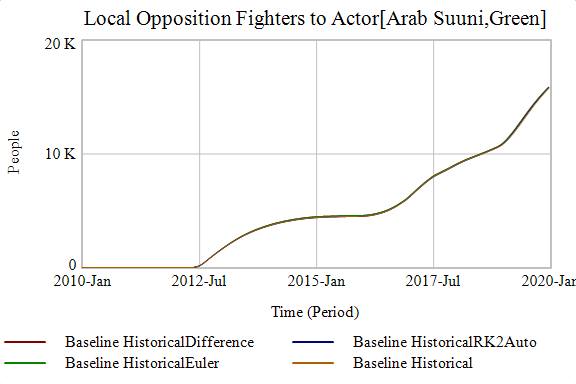
This example is just one way in which the model demonstrates robustness to extreme conditions. Performing in realistic ways after accepting a fantastic premise that one billion *Combatants* can be added. If they cannot be paid for, there’s little military benefit even as the soldiers begin defecting and choosing different sides upon returning to civilian life. Some of these populations are conquered by Red, who does not have the manpower to garrison them and faces large internal uprisings. Even when *Combatants* can be paid for – the military success does not eliminate the consequences of having to deal with a huge influx of people, which is an acceptable validation of the test.

# C-7: Integration Error

Earlier versions of E-SAM suffered from an integration error. This error most commonly occurred in parameters directly influenced by the time length of a battle, which originally was at a single day .011. This matched the DT interval of .011, violating the common rule of thumb that DT should be at least 1/4th-1/5th the smallest time interval in the model. This resulted in visibly observable as “choppy” behavior graphs in parameters such as *KIA per Million Population* which is directly tied to military losses in battles as well as terrorism deaths. Because *KIA per Million Population* influences through structure *Local Opposition Fighters to Actor* by creating an unstable environment that requires more garrison – this integration error did cause different behaviors to emerge over time. But improvements in structure and formulation (see System Tests below) have mostly eliminated this problem. In Figure XX and Figure XX Difference, Euler and RK2 Auto integration methods are compared against the *Historical Baseline* which uses an RK4 Fixed integration method for both *KIA per Million Population* and *Local Opposition Fighters to Actor.*



The behavior is still very choppy because battles are still short relative to a period, 3 days (or .033) rather than 1 day (.011) long. However this system choppiness does not propogate through the model as it used to as perceptions are smoothed in *Actual Garrison* structure. Which is why the behavior of *Local Opposition Fighters to Actor* is smoothed in Figure XX.



Examining in detail the numerical differences of integration reveal some slight differences between the four methods as depicted in Figure XX and Figure XX below. However these are not significant differences.

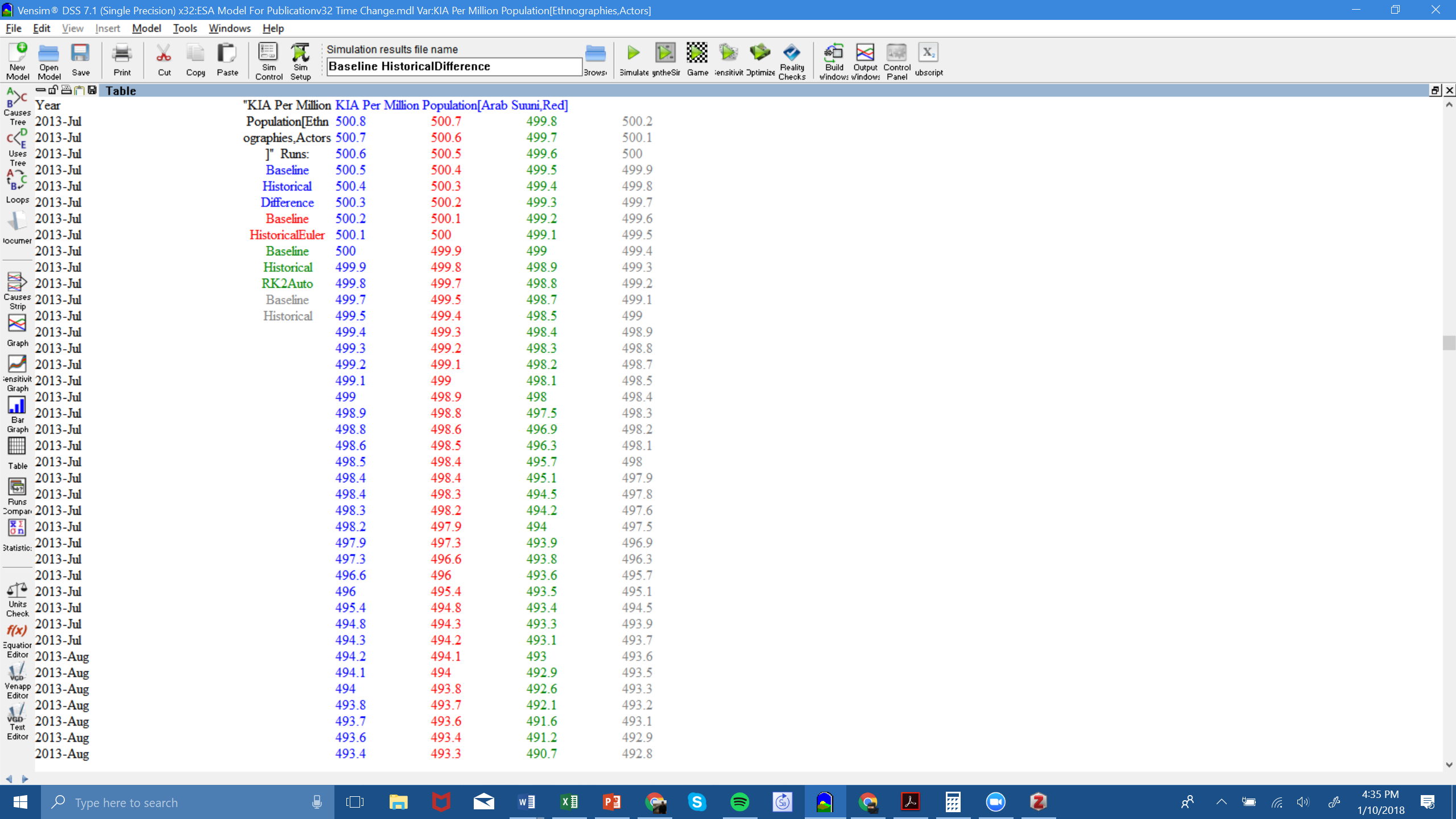


Figure : Integration Test - Differences in KIA per M by Method

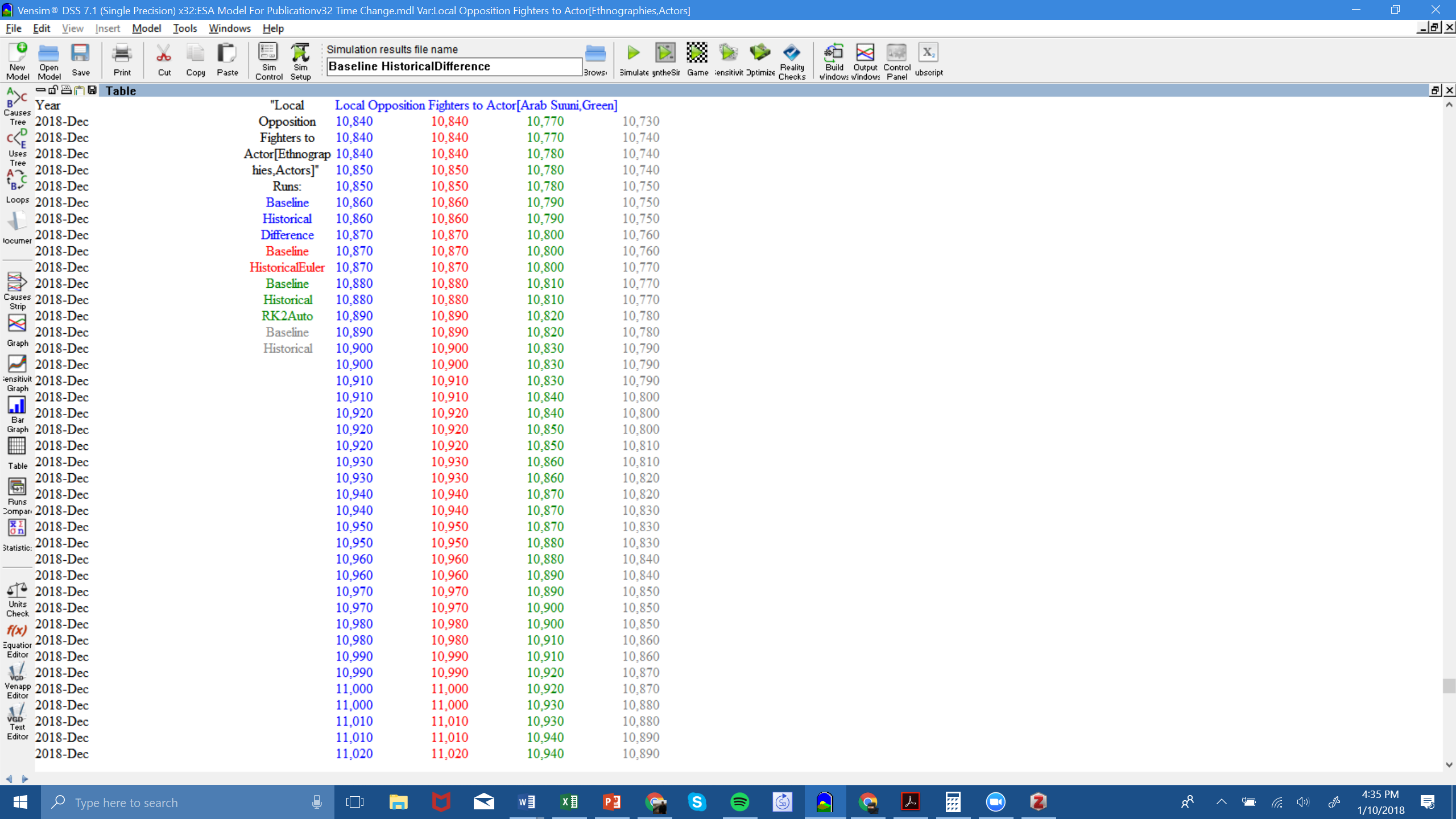


Figure : Integration Test - Differences in Local Opposition to Actor by Method

In some extreme cases integration error behavior may reappear. For example in the *LocalOnly* boundary test of removing foreign interventions, integration-error behavior appears again in the *Total KIA per Million Population* as seen in Figure XX*.* However, this choppiness does not propogate through to have an impact on the model as can be seen in comparing *LocalOnly* and *Historical Baseline* through *Local Opposition Fighters* in Figure XX.

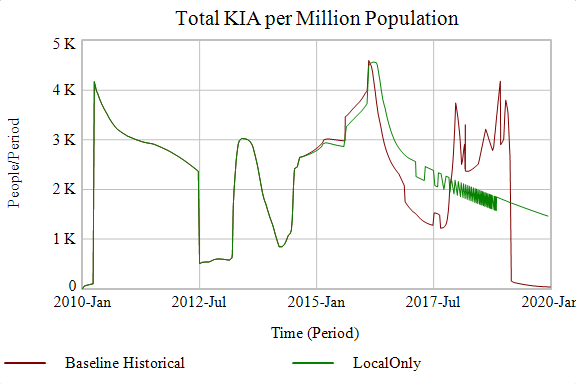


Figure : Integration Test - Integration Error Visible under Certain Conditions

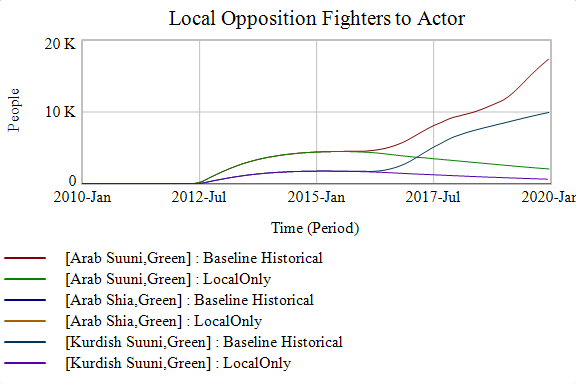


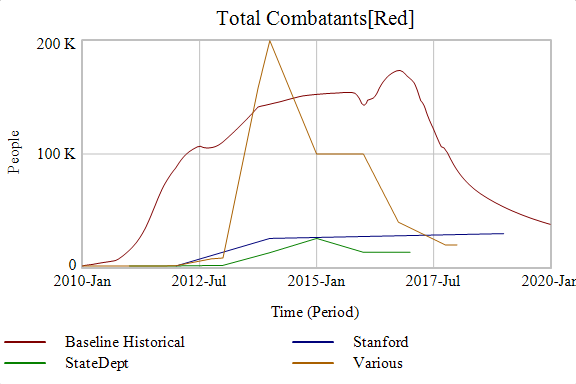
Figure : Integration Test - Structural Fix to Integration Error Demonstrated

Another example of where the choppy behavior appears in limited circumstances is in the *1 Billion Green* test where *Total Defections* showed choppiness in Figure XX above. This behavior however does not appear under normal circumstances and has limited effect on model performance.

# C-8: Behavior Reproduction

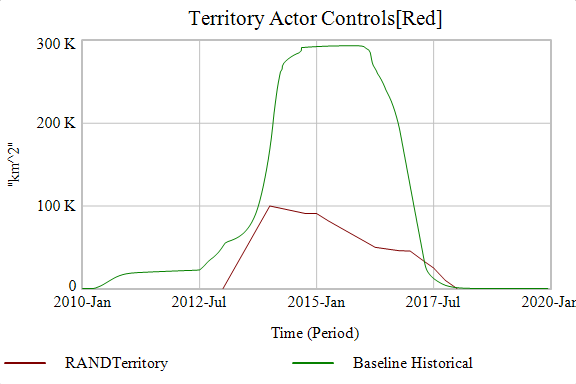
A challenge of building confidence through behavior reproduction is the paucity of data about ISIS itself to compare against. Although ISIS did publish some data on its own size and performance, this was clearly propaganda. Independent estimates, though many in number, typically provide a point in time and no behavior over time. By and large the most effective behavioral reproduction tests will therefore come from the reader comparing what they know and understand occurred to the behaviors in the model. Is it realistic and within reason to what they know to have occurred?

Three behaviors are reproduced related to *Total Combatants[Red]*. The first behavior is reproduced entirely from US State Department estimates. A second behavior comes from the Stanford University project Mapping Militant Organizations. And the third behavior is an composite of disparate reports on the size. These are compared against the Baseline Historical run in Figure XX.

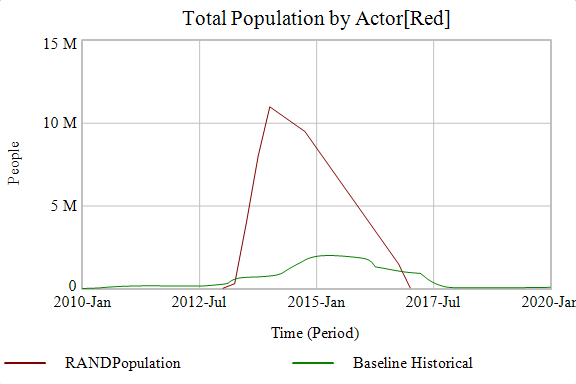


The Baseline Historical performance falls within the range of the low end established by the State Department and the Stanford effort, and underneath the high end of the various sources. It also roughly matches the behavior pattern of the US State Department and Various sources, in that there is a sharp rise, followed by a fall.

Size estimates of the total territory controlled, and the population falling within that territory by ISIS are even more difficult to come by than combatant estimates. RAND published a report claiming a certain peak level of ISIS control and subsequent declines by 2017. Interpolating behavior from those peaks and falls allows reproduction of gross behavior mode of *Territory Actor Controls[Red]* and *Total Population[Red]* to compare to the Historical Baseline in Figures XX and XX.



Although the magnitude is clearly off, the behavior closely replicates inflection points signaling the growth and decline of the territory held.



Interestingly the RAND population estimates are an order of magnitude larger than the model demonstrates, even though the territory controlled was an order of magnitude lower. This represents the difficulty in comparing datasets. However, the inflection points, and general behavior of rise and decline, remain consistent.

In summary although behavior reproduction is not a perfect match, often by magnitude in terms of peak, the key behaviors of growth and decline, as well as the timing of inflection points is presented as sufficient given the data limitations on the subject. As future peer-reviewed publications are made available with more consistent reporting of ISIS indicators over time these can be compared against the model results to continue to evaluate behavior reproduction.

# C-9: Behavior Anomaly

Behavior anomaly tests are a supplement to behavior reproduction tests when statistical comparison cannot be as easily established. This is more important in E-SAM because of its focus on realism versus precision and thus having a larger statistical error implicitly than a model that may have been fit through numerical calibration to a single specific historic behavior mode.

### Loop Knockout Tests

Loop knockout tests help established the importance of specific feedback structures by eliminating them from the model and seeing if performance significantly changes. Several of these have already been demonstrated in validation tests found elsewhere in this section:

* *Leadership Perception of Conflict Momentum:* When the feedback structure that connects *Territory Controlled by Actor* to *Offensive Stance based on Actor Perception of Momentum* is removed or simplified, unrealistic behavior emerges (see Figure XX in Structure Assessments section.)
* *Ethnographic Perception of Actors:* If feedback is eliminated between *Ethnographic Perception of Actor*  and *Rate of Unaligned Converting to Calculated Legitimacy* then Unaligned Population will fail to choose sides over time, resulting in a gradual accumulation of unreasonably high unaligned populations because they cannot “sense” what should be an obvious choice of actors. Note this isn’t the same as neither Actor being a good choice, which can occur in certain circumstances.
* *Fiscal Connections to Population and Territory:* Several feedback loops exist between the ability to gain revenue, pay expenses, and continue to do more actions. As was demonstrated in Figure XX when Extreme Conditions resulted in more troops than could be paid for – military activity ground to a halt. Likewise if the feedback between the various types of population controlled by an actor (Coerced, Calculated Legitimacy & Governed) and the ability to gain revenue off of them is severed – then the acquisition of population does not result in continued military capability, which is not realistic. This can be tested by setting *Criminal Activities per Person, Tax Rates* and *Territory Conditions Price per Unit* all to zero, in effect severing this feedback. The results of this test are examined in the primary measures of effect of *Territory Controlled by Actor [Red]* and *Total Combatants [Red]*  in Figures XX and XX. For this test all conditions are held the same as *Baseline Without Intervention* other than the loop knockout.

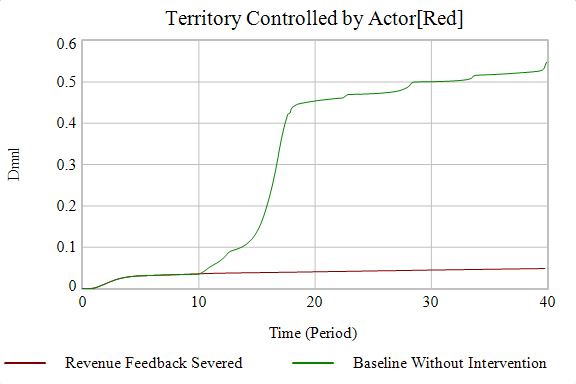


Figure : Loop Knockout of Revenue Feedback effect on Territory Controlled

As seen in the above figure, Red actor gains a minor amount of Territory at the periphery of Green but is not able to expand. This is because Red starts in the scenario with $10M USD. And initially *Total Combatants* begins rising due to the systemic repression of Green Actor against Arab Suuni and Kurdish Suuni pouplations. But without the feedback of finances, these *Combatants* cannot be paid and gradually winnow away through defections and other losses.

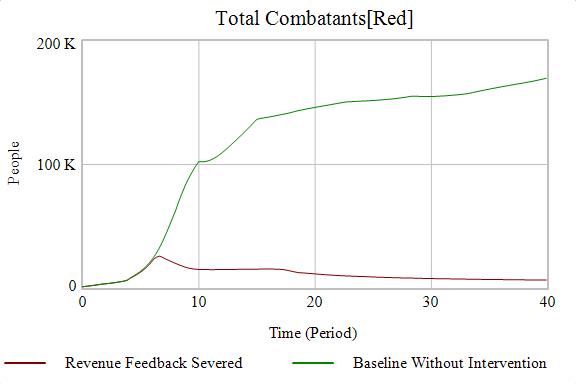


Figure : Loop Knockout of Revenue Feedback effect on Total Combatants

### Disequilibrium Tests

Tests that remove time delays involved in decision making – allowing instantaneous recognition and reaction to changing phenomena, can also identify behavior anomaly. For example removing all time delays related to ethnographic side-choosing results in anomalous behavior. This can be tested by reducing *Normal Time to Form Generational Perception* and *Time to Form Perception* in the Ethnographic Perception sector and *Normal Time for Population to Transition* in the Actor Legitimacy & Side Choosing sector (both part of World Model) to .01. This creates a near instant ability of an ethnographic group to alter there perception, and adjust between the three stages of legitimacy (up and down) of *Coerced, Calculated Legitimacy* and *Governed.* When these changes are made but everything else held constant to *Baseline without Intervention* primary measures of effect of *Territory Controlled by Actor* and XX can be compared against both baseline scenarios.

# C-10: Family Member Test

Since the E-SAM model is designed to be used in a wide variety of historical and regional circumstances to represent many forms of conflict less-than-full-spectrum, family tests are a key validation. Although the size of the model prohibits a full family test in this appendix, a single rudimentary family test can be constructed. Importantly the rudimentary family test should be fundamentally different than ISIS in Syria and Iraq which resulted in a full-scale insurgency and emerging-state actors. So recreating the Taliban in Afghanistan, Boko Haram in Nigeria would not be a good test.

### Family Test: Indonesia Counter-Terrorism Scenario

Instead a family-test will be created to explore capabilities in an environment where local fighters have not progressed outside of clandestine terrorist operations only. And the research question will be one being asked in many countries: given a small local ISIS population operating clandestinely and in jail, what happens when expatriate local fighters operating in Syria and Iraq on behalf of ISIS return?

To establish this scenario some liberties are taken. As ISIS is still the Red Actor, most parameters relating to ISIS performance will be held the same. Additionally even though the Indonesia Army has significant differences than Iraq and Syria, for purposes of a rudimentary family test most values will be held constant. What changes will be made are about ethnographic distribution, starting force size changes, and the addition of fighters abroad that can return to Indonesia who are not themselves foreign recruits.

To preserve subscript numbering for convenience, five of Indonesia’s many ethnographic divisions will be compiled into three groups: Javanese (43%), Sundanese (15%) and Malay, Madurese and Batak (10%). Although these three ethnographic groups only comprise 68% of the total 140M population normally they can be assigned proportionally equal waits to create a “whole” population (thus for now assuming there are no other ethnographies.)

Unlike the *Baseline* scenarios, in the *Indonesia Family* tests 100% political legitimacy at start will not be assumed. Rather the historical realities of previous conflict between civilian and government until 2006 will be taken into account, and 24% of the population will be set at *Calculated Legitimacy.* The military also has greater professional ability having not been devastated by invasion, civil war and sectarian divisions as was the case in Iraq & Syria. The scenario will start in the beginning of 2017. At this point ISIS in Iraq had largely been pushed back into Mosul, which was already under siege, and was losing ground steadily in Syria to a combination of Turkish, SDF supported by the US and Syrian forces.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Javanese | Sundanese | Malay |
| *Starting Ethnographic Population* | 88,014,450 | 31,800,766 | 20,636,448 |
| *Starting Population that is Calculated Legitimacy [Green]* | 24% | 24% | 24% |
| Starting Population that is Governed[Green] | 74% | 66% | 66% |
| Starting Ethnographic Generational Perception[Green] | 59,240,000 | 21,540,000 | 14,140,000 |
| Starting Ethnographic Current Perception [Green] | 31,700,000 | 11,400,000 | 7000000 |
| Starting Ethnographic Generational Perception[Red] | 3,500,000 | 1,300,000 | 825,000 |
| Starting Ethnographic Current Perception [Red] | 15,800,000 | 5,700,000 | 3,700,000 |
| Actual Desire to Credibly Govern[Green] | 76% | 76% | 76% |
|  |  |  |  |

|  |  |  |
| --- | --- | --- |
| Parameter | Green Actor | Red Actor |
| Starting Cash | $25B USD | $10M USD |
| Starting Total Combatants | 273,824 | 500 |
| Starting Detainees by Actor | 0.00E+00 | 270 |
| Starting Experience | 3 | 3 |
| Starting Worldwide Population of Foreign Recruits | 0 | 10,000 |
| Starting Combatants by Ethnography | 171,593  61,998  40,232 | 313  113  74 |
| Normal Experience Gained per Person | 1 | .5 |
| Starting AFV/IFV | 2137 | 0 |
| Starting Artillery | 594 | 0 |
| Starting Actor Conditions Expatriate Fighters | 0 | 300 |
| Starting Actor Security Effectiveness | 1 | .5 |
| Minimum Force Size to Engage | 0 | 20,000 |

The game script (see section B-6 Scenario Creation & Game Scripts) is:

:Time=0

STARTING ETHNOGRAPHIC PERCEPTION OF ACTOR

STARTING ETHNOGRAPHIC PERCEPTION OF ACTOR[Javanese,Green]= 88014451

STARTING ETHNOGRAPHIC PERCEPTION OF ACTOR[Sundanese,Green]= 31800766

STARTING ETHNOGRAPHIC PERCEPTION OF ACTOR[MalayMadureseBatak,Green]=20636448

STARTING ETHNOGRAPHIC GENERATIONAL PERCEPTION[Javanese,Green]= 65130694

STARTING ETHNOGRAPHIC GENERATIONAL PERCEPTION[Sundanese,Green]= 23532567

STARTING ETHNOGRAPHIC GENERATIONAL PERCEPTION[MalayMadureseBatak,Green]= 15270972

Actual Desire to Credibly Govern[Javanese,Green]=0.74

Actual Desire to Credibly Govern [Sundanese,Green]=0.74

Actual Desire to Credibly Govern [MalayMadureseBatak,Green]=0.74

Actual Desire to Credibly Govern[Javanese,Red]=0.0

Actual Desire to Credibly Govern [Sundanese,Red]=0.0

Actual Desire to Credibly Govern [MalayMadureseBatak,Red]=0.0

OpOrder Armed Civil Affairs[Green]=0.0

OpOrder Combatting Terrorism[Green]=0.6

OpOrder Conventional Warfare[Green]=0.2

OpOrder Indirect IED VBIED or SVIED[Green]=0

OpOrder Prison Breaks[Green]=0

OpOrder Prison Duty[Green]=0.0

OpOrder Propoganda[Green]=0.0

OpOrder Armed Civil Affairs[Red]=0.25

OpOrder Combatting Terrorism[Red]=0

OpOrder Conventional Warfare[Red]=0.2

OpOrder Indirect IED VBIED or SVIED[Red]=0

OpOrder Prison Breaks[Red]=0.05

OpOrder Prison Duty[Red]=0

OpOrder Propoganda[Red]=0.25

OpOrder Recruiting[Javanese,Green]=0.03

OpOrder Recruiting[Sundanese,Green]=0.01

OpOrder Recruiting[MalayMadureseBatak,Green]=0.01

OpOrder Recruiting[Javanese,Red]=0.01

OpOrder Recruiting[Sundanese,Red]=0.01

OpOrder Recruiting[MalayMadureseBatak,Red]=0.01

OpOrder Terrorism[Javanese,Red]=0.05

OpOrder Terrorism[Sundanese,Red]=0.05

OpOrder Terrorism[MalayMadureseBatak,Red]=0.05

:Time=39.99

In plain language what these settings and script represent is a very small indigenous ISIS presence in Indonesia. The group is militant – mixing terrorism, propaganda and armed civil affairs while seeking to recruit and grow their strength. They seek to target local prisons to free other combatants and will receive back fighters returning from abroad. They are not afraid of attacking the military conventionally – but will wait until their strength is at least 20,000 before launching an insurgency.

### Family Test: Indonesia CoutnerTerrorism Baseline

The baseline run of the above scenario results in a wholly different outcome than ISIS in Syria & Iraq. Red begins the scenario with the return of expatriated Indonesian fighters returning from Syria & Iraq (as seen in Figure XX) and prison breaks to free fellow combatants from Indonesia’s weakly secured prison system (see Figure XX).

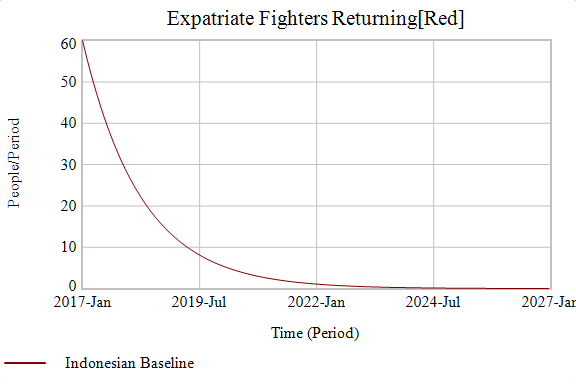


Figure : Family Test - Indonesian Fighters Returning from Syria & Iraq

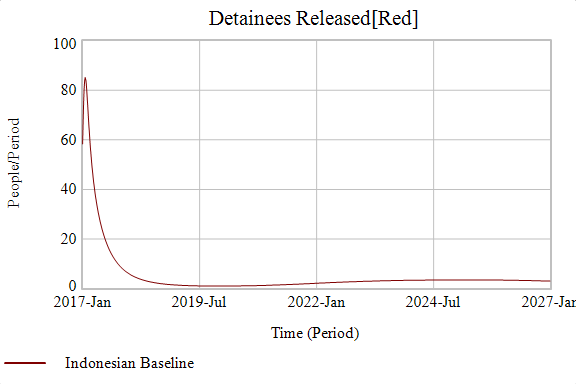


Figure : Family Test - Indonesian ISIS Fighters Released in Jail Breaks

This provides an initial increase in *Combatants* which allows Red as shown in Figure XX.

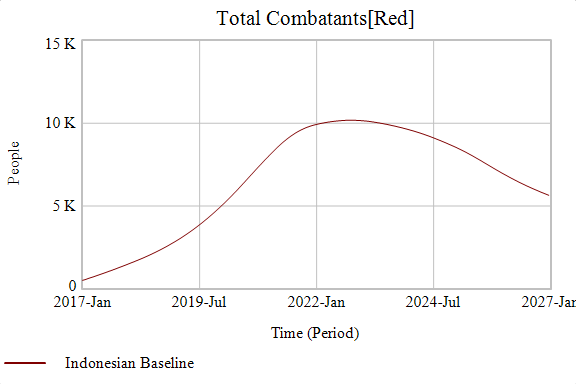


Figure : Family Test - Indonesia Growth of ISIS

As Red gains forces, it continues to wage a largely clandestine terror campaign against Green. This draws foreign fighters via propaganda into the Indonesia region to join Red as a result of their terror successes. The results are shown in Figure XX as both the Total Terrorist Attacks (right side vertical axis) and *Foreign Combatants* (left side vertical axis).

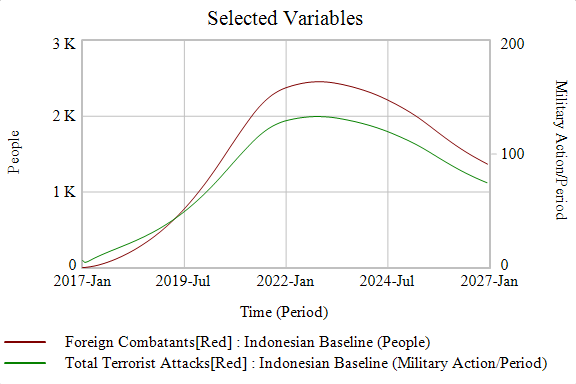


Figure : Family Test - CounterTerrorism Results for Green

Green is not idle during this campaign. It’s counter-terrorism forces are able to thwart attacks by Red. Red *Combatants* are killed in police raids while cells broken up result in more detainees of Red as shown in Figure XX.

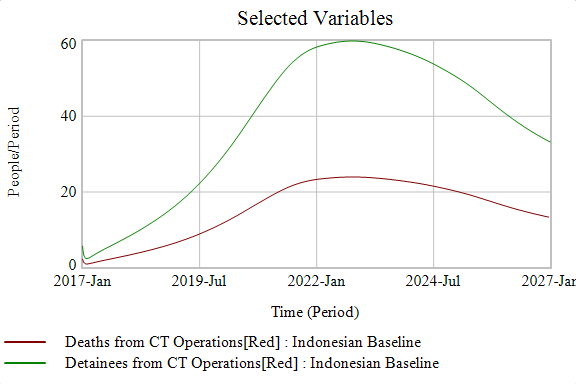


Figure : Family Test - Indonesia Baseline CT Results

Still civilians are dying from terrorist actions as demonstrated in Figure XX.

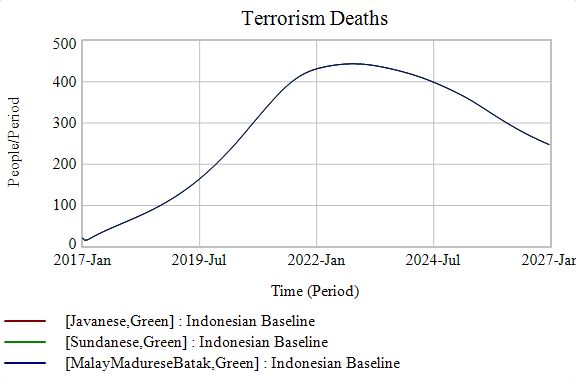


Figure : Family Test - Indonesia Terrorism Deaths

In terms of the narrative of legitimacy between the state and non-state actor, Green’s credibility is eroding over time as shown in Figure XX.

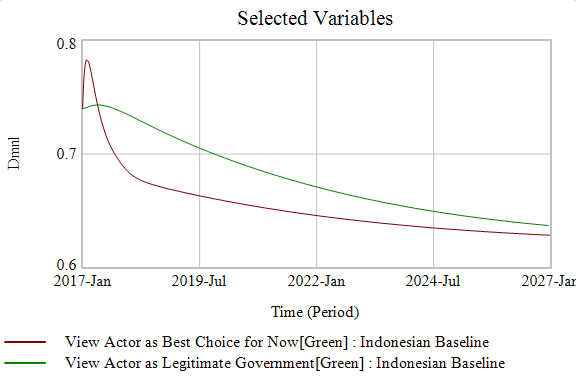
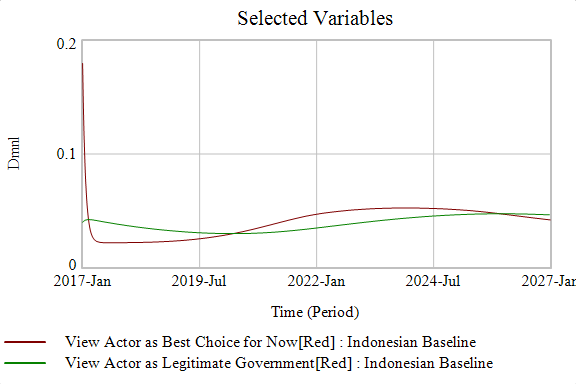


Figure : Family Test - Indonesian Legitimacy of Green

But Red, despite it’s propaganda and armed civil affairs, is unable to gain much support in the population.



The baseline demonstrates a persistent indigenous clandestine terrorism group causing harm – but not able to expand out of its niche. It reveals the weakness of Indonesian prisons which serve as a rotating training school. Combatants are captured by counter-terror operations, jailed where they gain more experience, and broken out by ISIS combatants later. The Red Actor is not able to launch an insurgency, let alone become an emerging-state actor. It is a threat best described as a law-enforcement and counter-terror operation than a military campaign. This is a realistic outcome, even if it is not precise.

But the proposed validity of E-SAM is not just in describing a hypothetical theater of operations – but allowing operational planners to prepare courses-of-actions (COA) for policy decision making. Suppose in this hypothetical the Indonesian government assigns operational planners a year in 2018 to prepare a set of policy recommendations that will be put in place to mitigate or contain ISIS from 2019 onward. To build confidence in this aspect of a Family Test three rudimentary COA’s are proposed to represent this exercise and the results compared to the baseline. These in effect are an additional form of validation as System Improvement Tests. Can reasonable policies that are actionable to decision makers result in noticeably different performance levels of the underlying system? Table XX below lays out the three COA’s including a plain language description and the specific game script changes that will be made at Period 8.01026 (e.g. Jan-2019).

|  |  |  |
| --- | --- | --- |
| COA | Plain Description of Strategy | OpOrders @ Period 8 |
| COA1 | Begin aggressive measures against ethnographies supporting Red. Reduced government services and tolerate sporadic incidents of extra-legal violence. | 1. Set Actual Desire to Credibly Govern [Green] to .5,.5,.5 2. Set OpOrder War Crimes [Green] to .001,.001,.001 |
| COA2 | Invite foreign intervention. Blue troops maintain non-combat capacity-building role supporting CT Training, Information Operations & Armed Civil Affairs on behalf of Green. | 1. Set Blue or Purple Intervention Size[Green] to 50,000 2. Set Information Operations, Counter-Terrorism Training of Green, Provision of Advanced Equipment and Armed Civil Affairs to .25 |
| COA3 | Isolate Red from ethnographies by increasing government services to ethnic groups while prioritizing Counter-Terrorism, Prison Security and messaging (Propaganda & Armed Civil Affairs) over military engagement. | 1. Set Conventional Military Actions[Green] to 0 2. Set Actual Desire to Credibly Govern [Green] to .8,.8,.8 3. Set OpOrder Combatting Terrorism [Green] to .8 4. Set OpOrder Prison Duty [Green] to .05 5. Set OpOrder Propoganda[Green] to .05 6. Set OpOrder ArmedCivilAffairs [Green] to .05 |

With these changes made in game scripts and run, the three COA’s are compared against the baseline in a variety of primary and secondary measures. A Family Test validation would be plausible behavior combined with surprising behavior from what are admittedly rudimentary scenarios.

First presented is the conflict narrative of how the population, across all Ethnographies *View Actors as Best Choice for Now* as shown in Figure XX (Green) and Figure XX (Red.) Although the amount who *View Actor as Legitimate Government* is also important in many factors, the calculated legitimacy level is useful for identifying how many are susceptible to switching.

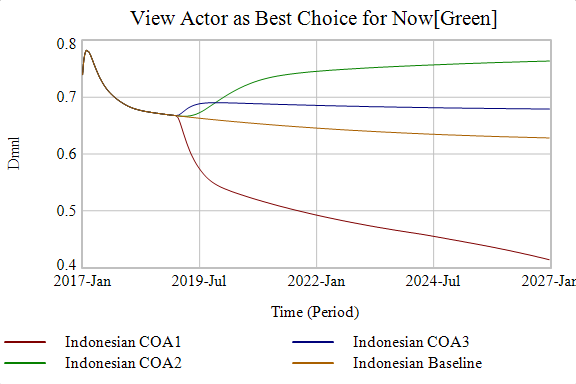


Figure : Family Test - COA Impact on Green Legitimacy

A surprising behavior emerges in Figure XX wherein although COA3 is specifically focused on improving governance, it is COA2 with a foreign intervention that results in more calculated legitimacy for Green. This will be explained further on an expanded section of Special Behavior in the Family Test.

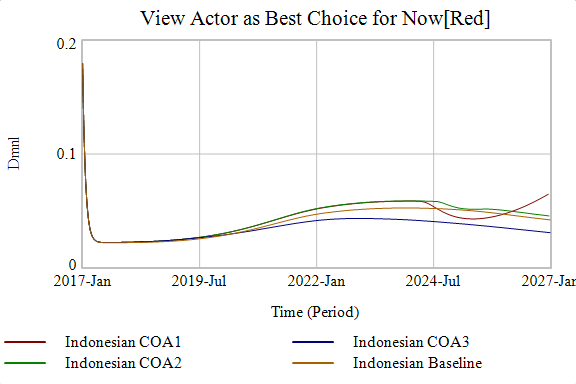


Figure : Family Test - COA Impact on Green Legitimacy

These two snapshots demonstrate the impact of each COA on the perception of Green. The divergence of perception created between COA1 (aggressive retaliation) and COA2 & COA3 is striking – indicating that although many do not support Red, they are losing support of Green.

The next primary measure of effect to examine is the level of *Total Combatants* (Figure XX).

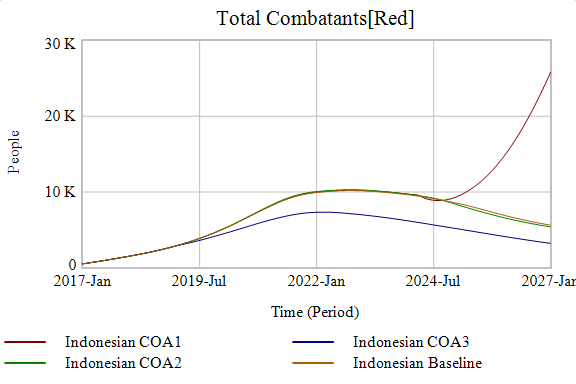


Figure : Family Test - COA Impact on Total Combatants

According to Kilcullen’s counterinsurgency theory of overreaction the results of COA1 are not surprising. Whereas COA3 has reduced the Red force to a fraction of its original size and COA2 at least maintains parity with baseline has advanced the system-state from clandestine terrorism to a full-blown insurgency. From the vignette settings, Red would not engage in *Conventional Warfare* until it had at least 10,000 *Combatants* which is achieved in COA1 near July of 2024. This results in beginning to fight for territory, as shown in Figure XX.

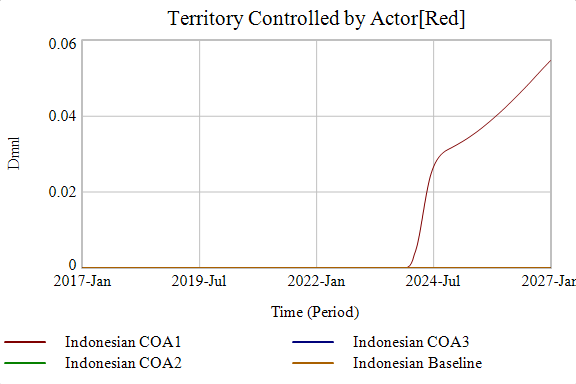


Figure : Family Test - COA Impact on Territory Controlled

In the Baseline run, as well as COA2 and COA3, Red Actor never felt sufficiently strong to begin launching territorial attacks. But in COA1 Red is able to begin taking territory, albeit at small levels with the following caveat. Since the territorial map of Indonesia was never established in the scenario initialization, and the location of geography, battle type and ethnographic distribution based on territory percentage was borrowed from the Iraq & Syria model – this may be an inaccurate representation of how much territory is captured when. However, the important result is that an insurgency did emerge, sufficient to begin *Conventional Warfare* and COA1 is the only scenario of the vignette where this happens.

A secondary measure of effect related to *Total Combatants* is the amount of ISIS fighters who were able to escape prison and rejoin the fight. This is shown below by COA in Figure XX.

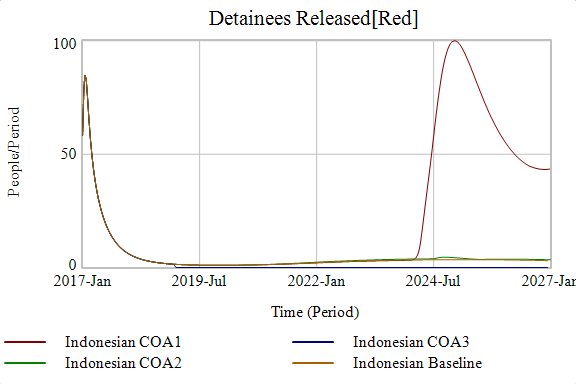


Figure : Family Test - COA's impact on Detainees Released

Not surprisingly COA3, which focuses on securing prisons shows the *Detainees Released* dropping to zero which is a crucial factor in reducing the size of *Total Combatants.*

Given the original vignette of a clandestine terror network operating in Indonesia, another primary measure of interest was the impact on the COA’s in terms of *Total Terrorist Attacks.*

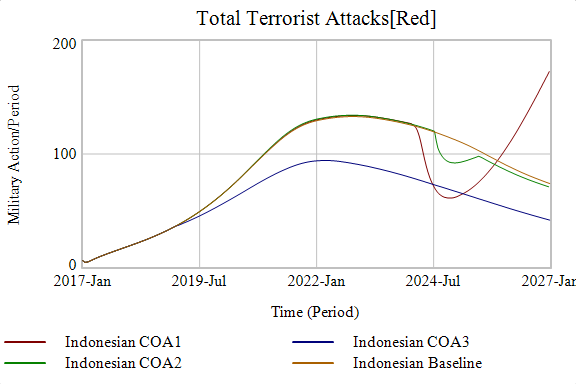


Figure : Family Test - Impact of COA's on Terrorist Attacks

Clearly in this comparison COA3 performs better than the other COA’s or baselines, effectively eliminating Red’s capability to conduct terror. As shown above COA3 had the lowest *Total Combatants* in large part by cutting off escaped combatants from prison and this correlates to overall terrorism levels.

COA1 again demonstrates the consequence of an overly militaristic and widely violent response in reaction to terrorism – resulting in more terror attacks over time. Another surprising behavior here however is that COA2, with foreign intervention aimed at improving security effectiveness of local troops, has not significantly diminished the terrorist attacks from the baseline. This will be covered in the surprising behavior section below.

Although more primary, secondary and tertiary measures of effect could be evaluated a general picture is beginning to emerge.

* COA1 with it’s focus on ethnographic retaliation, including extra-legal violence, performs worse than any other COA1. In fact only in COA1 can ISIS’s clandestine terror network transform into a full insurgency including conventional military attacks and seizure of territory.
* COA2 with a foreign intervention provides a mixed result. It provides higher legitimacy than the baseline, while failing to significantly improve on the baseline results of reducing *Total Combatants* and *Total Terrorist Attacks.*
* COA3 performs best by reducing the *Detainees Released from Jail* which reduces *Total Combatants* and virtually eliminates *Terrorist Attacks.* The government is viewed just as favorably as in COA2.

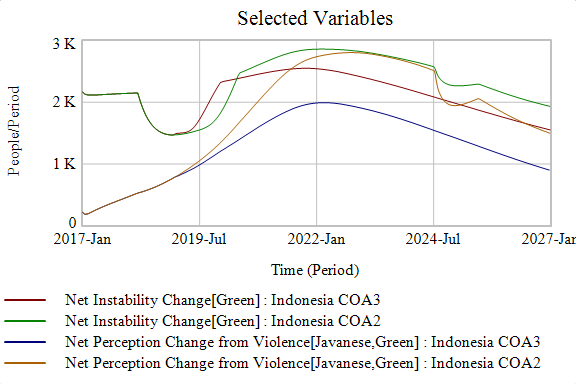
In summary although the Indonesian vignette is merely a sketch of a scenario with rudimentary courses-of-action compared, it still demonstrates a variety of plausible behavior for a low-conflict environment such as a clandestine terrorist-network threat. It can reproduce system behaviors as anticipated – not every conflict ends up in an insurgency or emerging-state actor; some “simmer” for years with low-popularity terrorist actors continuing to plague security forces while failing to gain their own foothold. It demonstrates the validation of being able to propose system improvement tests, and provide insights based on those tests, which can be used to inform decision making and policy creation. Finally even in this basic vignette E-SAM generated surprising behavior, as explained further below.

### Surprising Behavior: Indonesia Vignette

Part of the Family Test is to show that the model can be used within the same family on differing specifics. The generation of surprise behavior, which is a validation of the overall model – applies here as well. In this Vignette one notable surprise behavior occurred which is not observed in the *Baseline without Intervnetion* or *Baseline Historical* of the primary E-SAM model and is worth noting here.

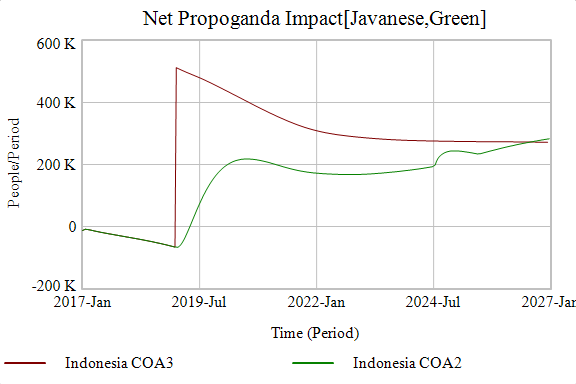
1. The first surprising behavior was when COA2, a foreign intervention supporting Green, resulted in higher favorability of the government than COA3 which focused on delivering more credible governance through Green.
2. The second surprising behavior is when COA2, which focuses on improving counter terrorism effectiveness, failed to significantly lower terrorism versus COA3.

The first surprising behavior results from the interaction of three contributors to ethnographic perception formation. The *Net Perception of Violence, Net Instability, Net Propaganda Impact,* and *Institutional Procedures* impact. In Figure XX the *Net Perception of Violence* and *Net Instability Change* are compared for COA2 and COA3.

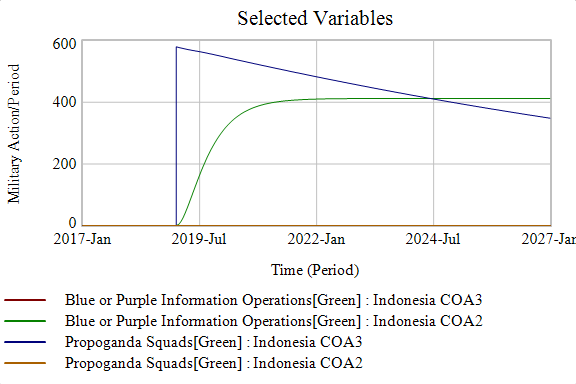


*Net Perception Change from Violence* is a factor of refugees from *War Crimes* committed by Green and refugees from *Terrorism* committed by Red plus the *KIA per Million Population.* Since these deaths are considered “the fault” of Green either for committing or not preventing, they have a multiplier. *Net Instability* is simply the overall *Conflict Deaths.* This clearly shows that COA3 with it’s more successful reduction in terrorism results in less ethnographic perception loss due to these factors of instability. But it’s only a measured in terms of thousands since the violence level was not that high to begin with.

Likewise in Figure XX *Net Propoganda Impact* is compared between COA2 and COA3.

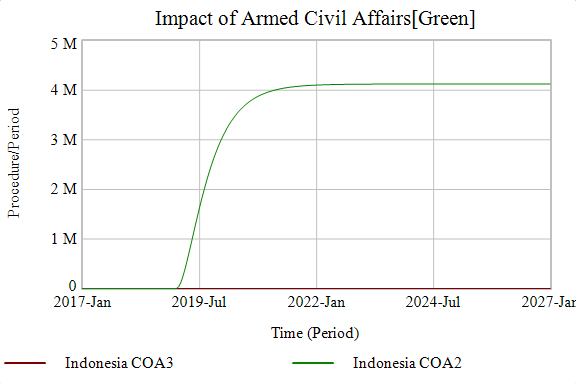


Again, COA3 is performing better in *Net Propoganda Impact* to begin with though the benefit of this gradually declines. This is because Green *Total Combatants* are constantly reducing due to *Defections*, which affects the allocation of Green in Propoganda – while foreign intervention forces remain constant. This effect can be seen in Figure XX.

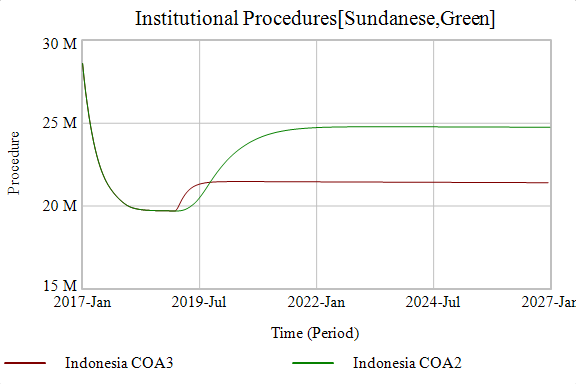


One benefit of foreign intervention is that the troops are less subject to local conditions of defections and thus may provide a more consistent basis of support.

But so far COA3 has generated slightly more positive *Ethnographic Perception* than COA2. So why does COA2 generate more legitimacy? It has to do with the credible number of *Institutional Procedures*, and specifically, the assignment of foreign troops to provide *Armed Civil Affairs.* These services provide a boost to credible government services where there might not otherwise be any. This provision of services is shown in Figure XX.



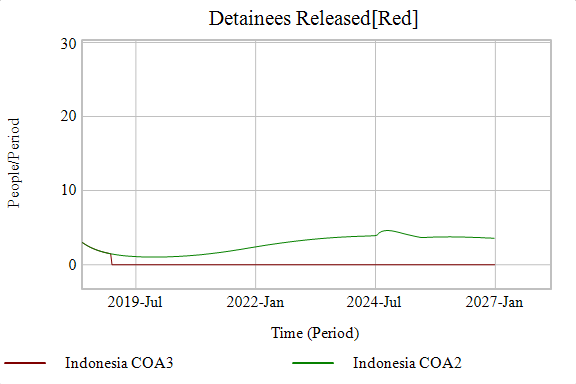
The benefit of these *Armed Civil Affairs* can be seen when looking at the Sundanese Ethnography *Institutional Procedures* in Figure XX.



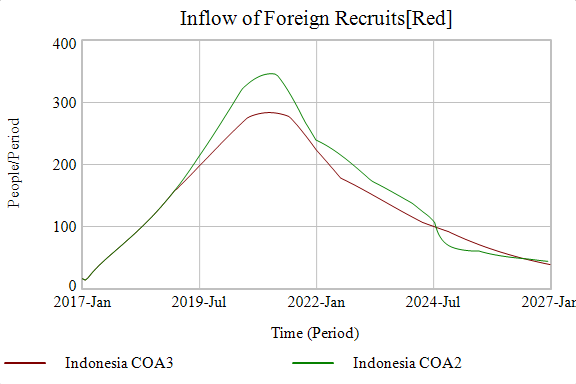
COA2 delivers nearly 3.5M more credible institutional procedural services, via the foreign troops providing civil affairs, than even the higher desire to provide governance in COA3. Whereas the violence and instability reductions experienced by COA3 only impact a few thousand people – these extra services help millions. It might be different in an area of high conflict or instability. So an important insight from this surprising behavior is that when violence is already low – provision of extra services may provide extra benefit because it will be realized by a larger percentage of the population.

The second surprising behavior of the Family Test is that COA2, despite providing foreign troops to increase the security effectiveness of Green’s counter-terrorism forces, arises from the “Revolving Door” archetype (see Sector-by-Sector Overview B-3 for Prison Breaks and Prison Duty.)

COA2 provided additional training to counter-terrorism forces, but left them at the same allocation as the baseline. COA3 not only increased that allocation by decreasing the use of conventional warfare, but also provided a higher priority to guarding prisons. This in effect cut the link of the “Revolving Door” archetype by ensuring that combatants once captured, remained in prison rather than breaking out to rejoin Red. The difference in this regard between COA2 and COA3 can be seen in the *Detainees Released* as shown in Figure XX, which is zoomed in on the implementation time of the COA to more clearly show the difference.



It doesn’t look like much – indeed only handful of extra combatants per period. But the “Revolving Door Archetype” includes the positive feedback loop of the more *Total Combatants,* the more *Terrorist Attacks* and the more *Terrorist Attacks* the more *Foreign Combatants.* So that handful of combatants adds up when a few more successful *Terror Attacks* take place and draw in several hundred more recruits as seen in Figure XX.



This “Revolving Door” effect undermines COA2’s increased security training at a critical juncture.

# C-11: Surprise Behavior

Over the course of model development many surprise behaviors have been identified. This is unsurprising given the limited amount of robust simulation modeling that had been performed on insurgencies and emerging-state actors prior. Much of this surprise behavior had to do with identifying certain timing windows of effectiveness. As emerging-state actors, as proposed by this theory and simulation, are path-depenednet actors early interventions tend to have

* Attacks against resource production are only truly effective befeore a population has given legitimacy
* Attacking governing capacity may be more effective in the later stage
* Allowing an emerging-state actor to expand past their ethnographic envelope may be advantageous strategically to ‘confining’ them within a friendlier population.
* Allowing Foreign Combatants to join the Red Actor actually in

# C-12: Sensitivity Analysis

Many of the parameters used in E-SAM will always be subject to high uncertainty due to subject matter. Non-state actors do not submit themselves to double-blind peer reviewed experimental studies while even state actors classify much of their information for security reasons. Each conflict contains unique local circumstances and unrepeatable sequences of events. General parameters may identify causal relationships, but the exact ratio of these relationships may never be known. Modeler judgement and expert insight must substitute when empirically observed data is not available.

This makes sensitivity analysis especially important to understand if values assigned to parameters are plausible. Sensitivity tests can also narrow down the field of potential parameters of interest to the true leverage points and thus help focus future research.

Traditionally sensitivity analysis covers numerical, behavioral and policy. Since the

purpose of E-SAM is to favor realism of causal relationships for research purposes and policy creation over point-behavior, numerical sensitivity (changes in magnitude) are not tested, and instead accepted to exist in the in the premise of the model. However, behavioral sensitivity (changes in the shape of the behavior mode) and policy sensitivity (“effective” policies become ineffective or worse on parameter change) are examined because these could have large ramifications. Research based on a sensitive parameters inaccurately represented may lead to incorrect findings. While policies based on the same could lead to unintended harmful actions or wasted resources. In terms of future studies identifying the highly sensitive parameters can provide guidance to future efforts to more precisely quantify these in different times, regions or circumstances.

The method of sensitivity analysis conducted was to identify thirty-six Starting Conditions, Constants and Time Delays. Each parameter was then examined individually for sensitivity in the following manner.

1. The model is set with starting conditions and all OpOrders follow the Historical Baseline run which incorporates historic interventions and replicates historical behavior.
2. Saved parameters to compare for sensitivity were selected from Primary Measures of Effect: *Territory Controlled[Red], Total Combatants[Red], and Total Population Controlled by Actor[Red.]*
3. Control parameters were selected to be tested individually, often by subscript.
4. A minimum and maximum value of the parameter was established.
5. Modify control parameter using Vensim sensitivity control modified across 200 runs with Latin Hypercube sampling along a normal random distribution (Noise Seed 1234).
6. Results were graphed using sensitivity strips in Vensim with bands at 50%, 75%, 95% and 100% of the runs. These graphs were then manually inspected and categorized as follows:
   1. Negligible Sensitivity: Zero to very little behavior pattern change across any band.
   2. Minimum Sensitivity: behavior pattern changes observed only in the 100% band range.
   3. Moderate Sensitivity: behavior pattern changes observed only in the 95% and 100% bands.
   4. Significant Sensitivity: behavior pattern changes observed in the 75% and above bands.
   5. High Sensitivity: behavior pattern changes observed in the 50% and above bands.

These ratings were assigned to both Behavioral and Policy sensitivity. Behavioral sensitivy focused on the inflection points, timing and magnitude of behavior compared to the Historical Baseline.

Policy sensitivity focused on fundamental changes of the shape of behavior that diverged from the Historical Baseline or even resulted in behavior that exceeded the counterfactual Historical Without Baseline behavior. Considering the runs were conducted with intervention, then a variable wherein change resulted in behavior by Red that exceeded even the case without intervention is of very high importance.

In each section below a summary table provides an overview of all parameters tested and the results. Then specific discussion follows on sensitive parameters and possible implications from a policy perspective.

### Starting Conditions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Starting Condition | Units | Normal | Minimum Value | Maximum Value | Behavior Sensitivity | Policy Sensitivity |
| Starting Ethnographic Generational Perception[Arab Suuni, Green] | People | 5,000,000 | 10,000,000 | 15,000,000 | High | High |
| Starting Combatants[Red] | People | 1,500 | 750 | 2,250 | High | Moderate |
| Starting Ethnographic Generational Perception[Arab Shia, Green] | People | 30,000,000 | 15,000,000 | 45,000,000 | High | Moderate |
| Starting WorldWide Population of Recruitable Actors[Red] | People | 500,000 | 25,000 | 75,000 | High | Minimal |
| Starting Actor Security Effectiveness[Green] | Dmnl | 0.5 | 0.25 | 0.75 | High | Minimal |
| Starting Detainees by Actor [Red] | People | 1,500 | 750 | 2,250 | High | Minimal |
| Starting Ethnographic Generational Perception [Arab Suuni, Red] | People | - | (2,500,000) | 2,500,000 | Significant | Minimal |
| Starting Ethnographic Generational Perception [Arab Shia, Red] | People | - | (2,500,000) | 2,500,000 | Minimal | Minimal |

Although the pool of *Starting Combatants, Starting WorldWide Population of Recruitable Actors* and *Starting Detainees by Actor*  all showed a high degree of behavioral sensitivity, this is unsurprising. All of these relate in one way or another to obtaining more combatants for Red earlier in the conflict, and thus can be expected to create sensitivity in the outcomes. However none of these resulted in outcomes that showed meaningful Policy Sensititivity except for *Starting Combatants* and that only for a handful of runs at the upper values.

The most interesting sensitivity results are starting *Ethnographic Generational Perceptions* to the various actors. Each Actor has a supporting or opposed ethnographic group – one which naturally favors them and one that dislikes them. In the *Baseline Historical* the Red actor enjoys more support, and thus higher starting perception, from *Arab Suuni* while the Green actor is supported by *Arab Shia.* Likewise, the Green actor is more opposed to *Arab Suuni* and Red Actor opposed to *Arab Shia.*

The sensitivity of *Ethnographic Generational Perceptions* varies based on this ethnographic relationship, whether they are supporting or opposed, and to which actor. The overview of these results are depicted in a below chart that matches Actors with their Supported or Opposing Ethnographic groups and the sensitivity results.

|  |  |  |
| --- | --- | --- |
| 7 | Supported By | Opposed To |
| Green | *Arab Shia*  High Behavioral Sensitivity  Moderate Policy Sensitivity | *Arab Suuni*  High Behavioral Sensitivity  High Policy Sensitivity |
| Red | *Arab Suuni*  Significant Behavioral Sensitivity  Minimal Policy Sensitivity | *Arab Shia*  Minimal Behavioral Sensitivity  Minimal Policy Sensititivy |

This overview implies from a policy space that the starting ethnographic perceptions matter more in relation to the Green Actor, than to the Red. And furthermore that it is the ethnographic perception of the group opposing the Green Actor that matters the most.

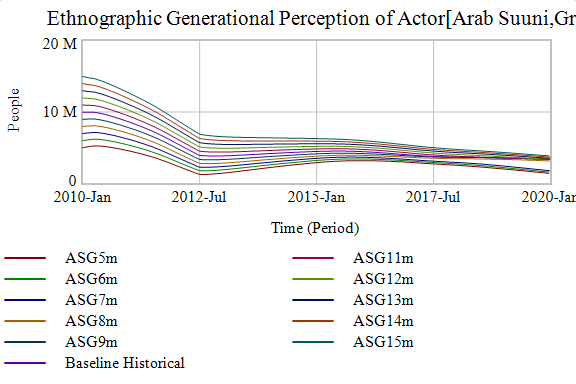
This can be demonstrated by comparing the sensitivity graphs of *Ethnographic Generational Perception [Arab Shia, Green]* and *[Arab Suuni, Green]* side by side below over the primary measures of effectiveness *Total Combatants[Red]* and *Total Population by Actor[Red].*

 **

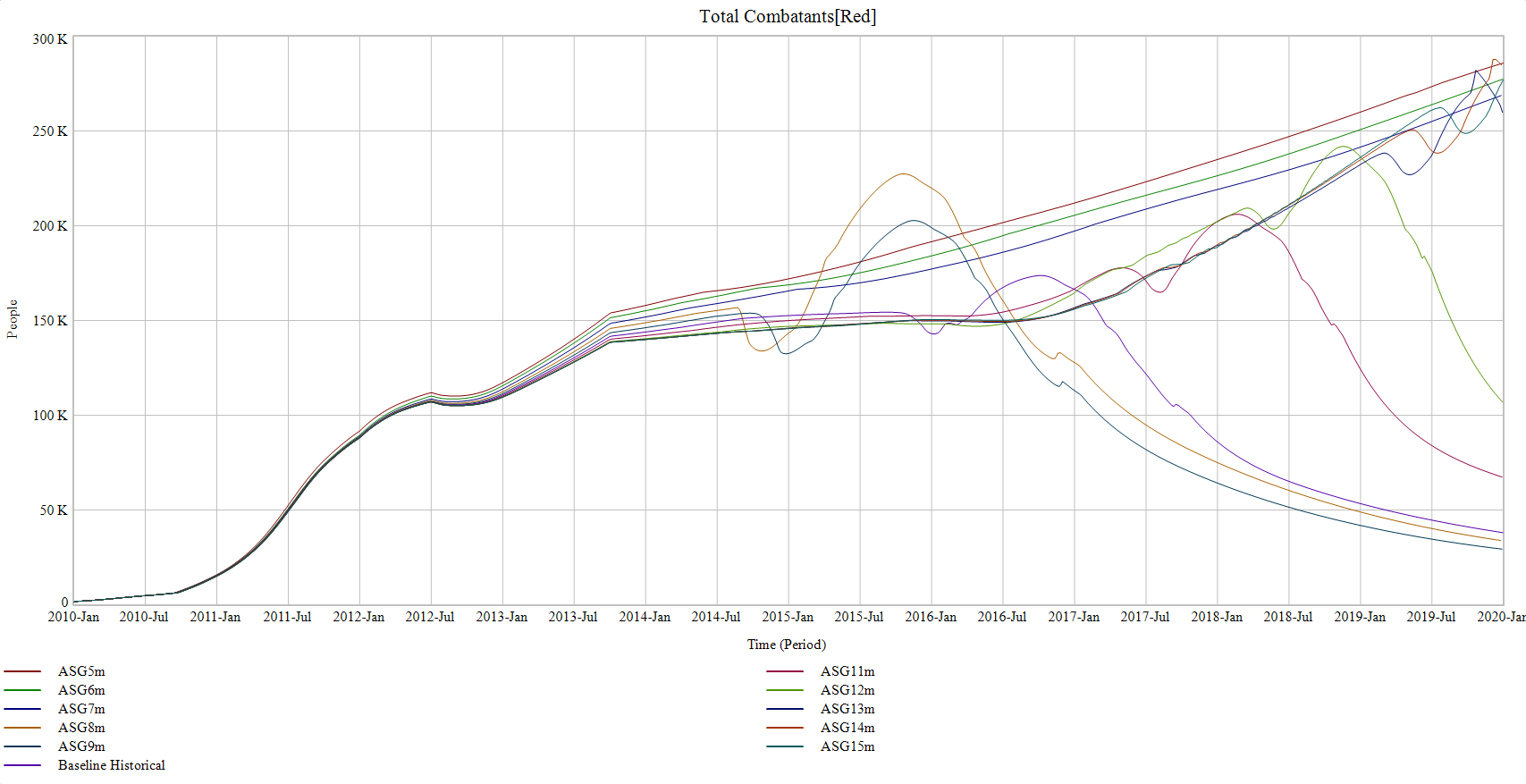
 **

To more closely examine this dynamic 10 runs, rather than 200, across the policy space from 5m starting ethnographic perception to 15m at increments of 1m is constructed. These are listed ASG5m through ASG15m. Note that the Historical Baseline begins with a normal value of 10m.

First a chart depicting the 15 runs shows that, as would be expected, when the opposition ethnographic perception starts higher, it ends higher.



However, the bifurcation of behavior around a hidden threshold becomes apparent when looking at *Total Combatants [Red]* as displayed in Figure XX.

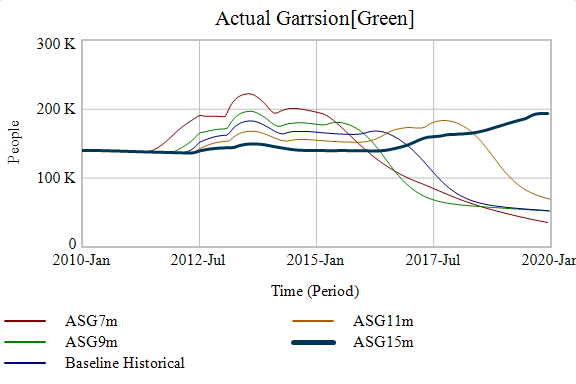


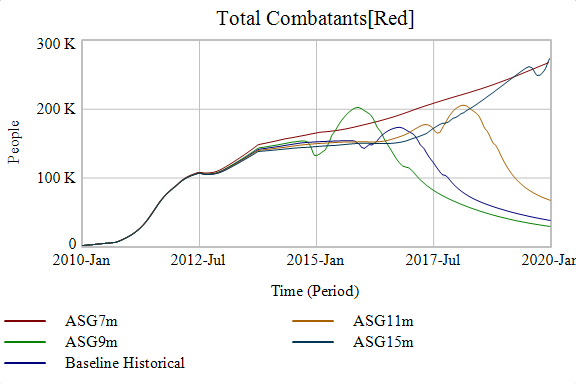
The runs split between growth for the Red Actor and collapse. But they split along two threshold points the lower value of which is between 7m-8m starting ethnographic perception and the higher occurs between 11-12m starting ethnographic perception. This is more clearly seen in the numerical table below of ending values for *Total Combatants[Red]*

|  |  |
| --- | --- |
| Run | Total Combatants[Red] Ending Value |
| ASG5m | 285,900 |
| ASG6m | 277,500 |
| ASG7m | 269,200 |
| ASG8m | 33,800 |
| ASG9m | 29,360 |
| Baseline Historical (10m) | 38,090 |
| ASG11m | 67,270 |
| ASG12m | 106,600 |
| ASG13m | 258,500 |
| ASG14m | 283,900 |
| ASG15m | 277,200 |

Note the first threshold point between ASG7m and ASG8m. This makes some sense – as the lower ethnographic support of ASG7m indicated Red Actor succeeded in mobilizing population grievances. The second threshold point between ASG11m and ASG12m is counterintuitive. Why would a very high starting ethnographic perception of an actor have a more-similar result to a much lower value than an in-between average?

This behavior, the second threshold, is caused because *Arab Suuni*, the opposing ethnographic group to the Green actor, has \*too\* favorable a position. As a result less military troops are required to garrison an unhappy population allowing Red Actor to expand more quickly when they begin. The *Actual Garrison[Green]* , displayed in Figure XX, is much lower for ASG15m than for ASG7m. ASG7m is the highest run that demonstrates success for Red Actor before dipping into the failure-valley. Likesiwe the *Total Combatants[Red]* at the same time period in Figure XX is higher.

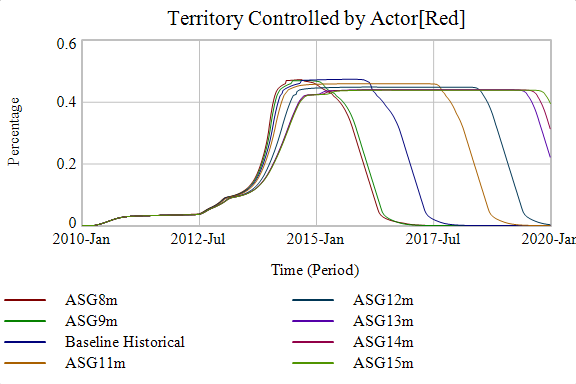




At ASG7m despite having a higher garrison the higher grievances of the population for *Arab Suuni* results in more *Total Combatnats[Red]* and Red is able to achieve a breakout.

ASG8m-ASG10m (Baseline Historical) however enter an ‘ethnographic-valley’. Arab Suuni’s grievances are supporting Red, but there is also a higher garrison and the tipping point is reached where the garrison is sufficiently large to slow Red’s expansion. Containing it to the point where the later foreign intervention can reverse the gains. The exit of that valley however pushes further to the right at ASG11m and ASG15m. A relatively calm population, with comparatively low garrisons stationed there, do not have the defenses in place to as quickly halt Red’s expansion.

It’s important to note that this second threshold is not a point of ultimate failure for Green as the first threshold point is. Looking at *Territory Controlled by Actor[Red]* in Figure XX below, any run with an *Starting Ethongraphic Generational Perception* at 8m or greater for *Arab Suunis* ultimately prevail against Red.



What changes is how long it takes to recover from the initial

The policy take-away from this sensitivity analysis is nuanced. It requires understanding whether there is an ethnographic population structurally opposed to the state power and then secondly understanding how their current beliefs can aid or hinder Red progress. Very low support leads to a Red success, while medium support combined with an adequate garrison leads to Red defeat. However, higher support with an ***inadequate*** garrison may represent a false-sense of security that doesn’t have the forces in place when Red attempts to break out.

### Constants

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Constants | Units | Normal | Minimum Value | Maximum Value | Behavior Sensitivity | Policy Sensitivity |
| T3R [Red] | Pct | 0.05 | 0.025 | 0.25 | High | High |
| T3R [Green] | Pct | 0.3 | 0.15 | 0.45 | High | Significant |
| Normal Experience Gain [Red] | Exp Years | 0.5 | 0.25 | 0.75 | High | Minimal |
| Normal Experience Gain [Green] | Exp Years | 0 | 0 | 0.5 | High | Negligible |
| Deaths per War Crime [Arab Suuni, Green] | People/ Military Action | 25 | 12.5 | 25 | Significant | Minimal |
| Advanced Equipment Modifier[Green] | Pct | 0.25 | 0.125 | 0.375 | Minimal | Minimal |
| Deployment Time[Green] | Period | 2 | 1 | 6 | Minimal | Negligible |
| Normal Recruiting [Red] | People/ Military Action | 10 | 5 | 15 | Minimal | Negligible |
| Average Experience of Local Recruit [Green] | Exp Years | 3 | 1.5 | 4.5 | Negligible | Negligible |
| Average Experience of Local Recruit [Red] | Exp Years | 3 | 1.5 | 4.5 | Negligible | Negligible |
| Blue or Puprple T3R [Green] | Pct | 0.67 | 0.52 | 0.82 | Negligible | Negligible |
| Deaths per War Crime [Arab Shia, Red] | People/ Military Action | 25 | 12.5 | 25 | Negligible | Negligible |

The highest sensitivity in Actor Constants has to do with the T3R rating for both Green and Red. This logistic ratio is a percentage that reduces *Total Comabants* down to an actual fighting force separate from logistics, administration, headquarters and other non-combat functions.

The three charts below demonstrate the high degree of sensitivity for primary measures of effectiveness for Green when controlling T3R.





Likewise, sensitivity is very high when T3R is modified for Red Actor.





What’s occurring here is not just an absolute, but a relative measure between the T3R’s of the competing actors. The closer they are to 1, equal T3R, the more *Total Combatants[Red]* there must be relative to Green. But the more lopsided a ratio, the more asymmetric Red becomes versus Green. Interestingly the T3R ratio of Blue Deployments in support of Green have far less sensitivity.

This policy implications of this sensitivity to the local actor’s T3R illustrates Kilcullen’s Article #22 “Local forces should mirror the enemy, not ourselves.” It represents a potential “fixes that fail” system archetype shown below.



The green balancing loop is the perception by Blue that the training level of Green is insufficient to halt Red. But the training goal is to make Green “look more like Blue.” But this time-delayed fix activates a vicious cycle, the red reinforcing loop, with an even longer time delay. Training Green in the image of Blue results in an increase in Green’s own T3R as the logistical, administrative, and headquarters operations begin to mimic Blue’s. This increase in T3R of Green relative to Red, increases the asymmetry of Red relative to Green. The asymmetry can be thought of the ratio of every 100 *Total Combatants* that each Actor possess: how many are conducting actual military actions?   
  
In the Baseline Historical for every 100 *Total Combatants* Red is able to use 95 of them to perform military actions. Green only is able to convert 70 of them into military actions. The resulting asymmetry is ~1.3:1 and, when combined with the unequal distribution of forces by Green relative to where Red is attacking, accounts for a large difference of the ability of Red to ‘punch above its weight.’ In the sensitivity analysis, the largest theoretical asymmetry is a ratio that occurs when Green is at the Maximum T3R of 45% and Red is at the Minimum of 2.5%. At this point the asymmetry in Military Actions per 100 Combatatants will be ~1.77:1 in foavr of Red. Although this doesn’t seem like much of an increase, it is sufficient to cause Red’s performance to exceed that of even the Baseline Without Intervention.

From a policy perspective this identifies two leverage points: minimizing the increase of asymmetry in T3R by following Kilcullen’s 22nd Article. A second leverage point is introducing an increase in T3R burden into a Red actor. This isn’t easy, as Red isn’t being trained by Blue – and may be an inadvertent 2nd order effect of some other action. But efforts designed to increase the logistical burden of Red, taking *Total Combatants* away from military actions and into supporting T3R, can help shift the asymmetry more favorably to Green.

### Time Delays

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Time Delays | Units | Normal | Minimum Value | Maximum Value | Behavior Sensitivity | Policy Sensitivity |
| Normal Time to Transition Population [Arab Suuni, Green] | Period | 0.25 | 0.125 | 0.375 | High | High |
| Normal Time to form Current Perception | Period | 0.5 | 0.25 | 4 | High | High |
| Normal Period | Period | 1 | 0.33 | 4 | High | High |
| Normal Time for Generational Perception to Form | Period | 10 | 5 | 15 | High | Significant |
| Normal Time to Transition Population [Arab Suuni, Red] | Period | 0.25 | 0.125 | 4 | High | Minimal |
| Normal Procedural Decay Fraction[Green] | Period | 5 | 2.5 | 10 | Significant | Negligible |
| Normal Procedural Development Time[Green] | Period | 2 | 1 | 4 | Significant | Negligible |
| Deployment Time | Period | 1.5 | 0.5 | 6 | Mimimal | Negligible |
| Average Time to Absorb Training[Green] | Period | 2 | 1 | 4 | Minimal | Negligible |
| Normal Procedural Development Time[Red] | Period | 2 | 1 | 4 | Minimal | Negligible |
| Time to Form Perception on Foreign Troops[Arab Suuni] | Period | 0.5 | 1 | 1.5 | Minimal | Negligible |
| Time to Form Perception on Foreign Troops[Arab Shia] | Period | 0.5 | 1 | 4 | Minimal | Negligible |
| Average Time for Anchor Security Effectiveness to Change[Green] | Period | 5 | 10 | 15 | Neglible | Negligible |
| Normal Time to Transition Population [Arab Shia, Green] | Period | 0.25 | 0.125 | 0.375 | Negligble | Negligible |
| Normal Time to Transition Population [Arab Shia, Red] |  | 0.25 | 0.125 | 0.375 | Negligble | Negligible |
| Normal Time for Unaligned to Choose Sides | Period | 10 | 5 | 15 | Negligble | Negligible |
| Organic Procedural Development Time[Red] | Period | 0.25 | 0.125 | 4 | Negligible | Negligible |

For most adjustments to Time Delay parameters the results are either a high degree of behavioral sensitivity, with no policy sensitivity, or negligible amounts of both. The overall shape of the behavior remains consistent, but the timing points of inflection, the specific magnitude and ending levels vary. This can represent an ambiguity in the knowledge about time delay affects that aren’t well studied. It may also be an area of modeler choice in adjusting these time delays to reflect different circumstances. Historical conflicts, or those in areas of low technological access might have longer time delays in perception formation because of how information travels more slowly. It may also reflect another kind of perception formulation difference where elites may still have access to more recent information, but the general population simply doesn’t have much more than word of mouth. In recreating historical scenarios adjusting these time delays can be used to represent such historical or circumstance specific conditions.

Side-Choosing

Aside from this general observation there is a specific category of sensitive parameters which are the three time delays that deal with ethnographic perception and transitioning of legitimacy between the two actors. *Normal Time for Generational Perception to Form*, *Normal time for Current Perception to Form* and *Normal Time to Transition Population* fall in this category. Taken together these three parameters are the cumulative time-delays of an information-flow that drives ethnographic side-choosing between the Green and Red Actor. However, the sensitivity does not hold true across all ethnographic groups. Table XX, similar to Table XX previously, maps the supporting and opposing ethnographic support to show how these relationships relate to sensitivity.

Table : Sensitivity of Ethnographic Perception & Transition Parameters

|  |  |  |
| --- | --- | --- |
| Actor // Ethnographic Stance | Supported By | Opposed To |
| Green | *Arab Shia*  Neglible Behaviroal Sensitivity  Negligible Policy Sensitivity | *Arab Suuni*  High Behavioral Sensitivity  High Policy Sensitivity |
| Red | *Arab Suuni*  High behavorial sensitivity  Minimal policy sensitivity | *Arab Shia*  Negligible Behavioral Sensitivity  Negligible Policy  Sensitivity |

The nature of the relationship between ethnographic group and actor as it relates to side-choosing plays a large role. Only Arab Suunis, who oppose and are opposed by the Green Actor and are favored by the Red Actor demonstrate this side-choosing time specific sensitivity. Arab Shia, when subjected to the same sensitivity on both Green and Red Actors shows negligible sensitivity.

This behavior is intuitively plausible. In the Baseline Historical scenario it is the Arab Suuni that Green targets for reduction in services and extra-legal violence that provokes the conditions of instability. From a Green perspective the information-flow the total time it takes the ethnographic group to perceive this mistreatment, adjusting both current and long term perceptions, then act on transitioning from from *Governed,* to *Calculated Legitimacy* and then finally to a *Coerced* state as they exit the sphere of influence of the state government. Along the way revenues and recruiting for Green will drop while the *Total Garrison Required* will rise, and if Green cannot meet that number *Local Opposition Fighters* will begin appearing. It is from these amorphous *Local Opposition Fighters* that Red gains some of its first recruits. And it is from the *Unaligned* population that side-choosing begins as local networks select Red over Green as being more aligned with their interests. From the Red perspective it is the speed with which Arab Suuni switch sides and grant legitimacy to Red as an alternative government that fuels many dynamics in the model such as recruiting, taxation and garrisoning levels.

Among the three parameters *Normal Time to Form Current Perception* has the most sensitivity, followed by *Normal time to Form Long Term Perception* and *Normal Time to Transition Population[Green]* with *Normal Time to Transition Population[Red]* creating the least sensitivity within this group*.* This is can be demonstrated by comparing sensitivity strips of *Total Combatants[Red]* across all three parameters.





In summary the sensitivity analysis served two purposes. It identified those parameters that could benefit from additional rigor in parameterization that may suggest future research opportunities. The second purpose was to identify areas of policy leverage that might not be immediately apparent. This most important aspect involves the side-choosing dynamics of ethnographic groups in relation to Green and Red. These side-choosing dynamics require a sophisticated understanding of an ethnography’s relationship with an actor (supporting or opposing) and how that relationship shapes simulation results. Side-choosing policies have implications not only for how to react to a conflict, but in shaping the conflict as well ahead of time and perhaps staving off a conflict in the first place. Another point of leverage is the logistical footprint, expressed in T3R ratios, of the Green Actor relative to Red.

Additional work still remains in creating multivariate sensitivity tests that could help understand E-SAM’s behavior in different environments. Numerical statistical analysis would also help understanding the relative strength of sensitivity. These remain for future efforts.

# C-13: System Improvement

System improvement is demonstrated when evidence can be collected and shown that an intervention proposed by a model resulted in the expected change. As the E-SAM has not yet been used in this manner – no such data can be collected and this remains an area for continued application and research. Such efforts should not just focus on whether the model behavior was realistic to the result of a policy – but also whether users increased their understanding before and after the use of E-SAM.