1818 Library ◆ Suite 600 ◆ Reston, VA 20190-5602 ◆ (703) 326-2861  
<http://www.metsci.com>



Simulator Cleanup

V1

June 4, 2013

Thomas M. Kratzke

Sr. Analyst

Contents

[1. Introduction 1](#_Toc358108259)

[2. Cleanup of the Input Files 2](#_Toc358108260)

[3. Cleanup of the Output Files 3](#_Toc358108261)

[3.1 Cleanup of the particle file 3](#_Toc358108262)

[3.1.1 Dimensions and Variables 3](#_Toc358108263)

[3.2 Cleanup of the Log File 4](#_Toc358108264)

[3.2.1 POS Values 5](#_Toc358108265)

[3.2.2 TIME tag (attributes of the TIME Tag) 5](#_Toc358108266)

[3.3 Cleanup of the Statistics File 6](#_Toc358108267)

List of Algorithms

List of Code

**No table of figures entries found.**

List of Figures

**No table of figures entries found.**

List of Formulas

**No table of figures entries found.**

List of Snippets

**No table of figures entries found.**

List of Tables

**No table of figures entries found.**

Document Revision History

|  |  |
| --- | --- |
| Version Date | Comments |
| May 20, 2013 | For *Sim* V1.5 |
| June 4, 2013 | For *Sim* V1.5 |
|  |  |
|  |  |
|  |  |

# Introduction

In this article, we present proposed cleanups of SAROPS simulator. We have divided the cleanup into two sections; the input and the output.

The input is essentially one file, called “the xml file.” Within that file, there are references to two environmental files, but those environmental files are not changing.

In the next section, we specify which parameters in the xml file are not used and could be deleted unless they are used by other programs. We also propose a structure that would simplify the parameter passing.

The section after that discusses Sim’s output files. There is a particle file and there is a “log file.” There is also a “stats,” file which we propose to eliminate.

# Cleanup of the Input Files

We refer to a tag’s “path” in a canonical way. For example, if the main tag is “SIM,” and it has a subtag “REQUEST,” which in turn has a subtag “CASE,” then the path is “SIM.REQUEST.CASE.”

There is little that can be omitted from the xml file until near the end. However, the tag SIM.REQUEST.CASE is completely unused by Sim and can be omitted.

Moreover, much could be simplified if we used certain conventions or defaults. Suppose a case is stored in a directory whose full path ends with <case-name>. Suppose we agreed that every Sim run had the name < case-name>-Sim.xml, and that every environmental file would be in < case-name>/Env. Moreover, every output file from Sim (or Planner) would go in < case-name>/Out, and the particle file name would always be < case-name>.nc. As an aside, we could also insist that every planner xml file be called < case-name>-Plan.xml.

Since Sim and Planner now keep track of the files they use in the directory < case-name>/EngineFiles, it would be much easier to read the directories and keep track of a case’s progress. Furthermore, it would eliminate the need for very lengthy file names in the ENVDATA.CURRENTS and ENVDATA.WINDS tags, and completely eliminate the need for the REQUEST.OUTPUT.file attribute. It would also make the xml more portable since everything would be in relation to < case-name> and the directory could be moved without updating the full path names typically embedded within the xml file.

Summarizing, the two cleanups for Sim.xml would be to impose the above structure on the input directory, and eliminate the tag SIM.REQUEST.OUTPUT.CASE from the xml file.

# Cleanup of the Output Files

## Cleanup of the particle file

The main output file that Sim produces is a NetCdf file called the Particle File. A NetCdf file can be thought of as a collection of multi-dimensional arrays, and here we describe these arrays.

### Dimensions and Variables

The arrays that we store are indexed by one or more sets, and NetCdf calls these sets *Dimensions*. The Dimensions that are of interest to us are “Scenario,” “Time,” and “Particle-index-within-scenario.” We abbreviate “Particle-index-within-scenario” as simply “Particle” in this section. We create a Dimension by specifying its name and number of elements.

In the table below, we list the variables that are in the current particle file. We propose to eliminate “Particle” and “Scenario,” and possibly “UnderwayTypes.”

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Dimensions | Type | Description |  |
| Time | Time | Int | Set of times for this Particle File |  |
| Particle | Particle | Int | Index within scenario |  |
| Scenario | Scenario | Int | Index within this case |  |
| ScenarioName | Scenario | Short | input id of scenario |  |
| ScenarioWeight | Scenario | Float | input weight of scenario |  |
| DistressTime | Scenario,Particle | Int | time at which this particle went into distress |  |
| DistressType | Scenario,Particle | Int | id of what the particle became at distress |  |
| UnderwayType |  | Scenario,Particle | int | id of what the particle started as |
| LandingTime | Scenario,Particle | Int | when it landed |  |
| Lat | Scenario,Particle,Time | Float | where the particle was at the given time |  |
| Lng | Scenario,Particle,Time | Float | where the particle was at the given time |  |
| Probability | Scenario,Particle,Time | Float | The proportion of weight that this particle was at this given time |  |
| PFail | Scenario,Particle,Time | Float | The cumulative pFail for this particle was at this given time |  |

In addition, there are attributes (strings) that are not related to any dimensions. These are essentially comments, but they can be read and we use them to store the model. In this way, we record the model that was used to create the Sim run.

The attributes we store are listed below. We propose eliminating xmlmodel2.

|  |  |  |
| --- | --- | --- |
| Variable | Attribute | Value |
| Time | Units | seconds since 01/01/00 |
| Xmlmodel | -- | The string containing the xml representation of the model |
| xmlmodel2 | -- | The string containing the xml representation of the model |

## Cleanup of the Log File

The log file is produced by the Simulator. This file contains the values of different forms of POS, as well as the “out-of-area” incidents.

This is an xml file that is primarily a list of time tags. Although only the last time tag is used, it is possibly useful to look at earlier time tags and so we are *not* recommending their elimination. We would like to examine which attributes are being used in these time tags.

### POS Values

A recurring theme is that we often are reporting probabilities as if there had been no search, and also reporting probabilities that take into account the search. If we update the weights of the particles by taking into account the search, we call this “*probability with search*,” and if do not update the weights, we call this “*probability without search*.”

### TIME tag (attributes of the TIME Tag)

For each time step, there is a “TIME” tag. Within the tag itself (and not the sub-tags), there are the following entries:

|  |  |
| --- | --- |
| distressParticlesWithSearch | *Probability with search* (at this time) of being in distress |
| landedParticlesNoSearch | *Probability without search* of being landed (and in distress) |
| landedParticlesWithSearch | *Probability with search* of being landed (and in distress) |

##### SCENARIO Tags

Within the “TIME” tag, there is a “SCENARIO” subtag for each scenario. A subtag of the “SCENARIO” subtag gives geographic information about the scenario. In the “SCENARIO” tag itself, there are the following entries:

|  |  |
| --- | --- |
| distressCount | Number of Particles in the scenario that are in distress |
| distressParticlesNoSearch | *Probability without search* of being in this scenario, given that it is in distress |
| distressParticlesWithSearch | *Probability with search* of being in this scenario, given that it is in distress |
| jointDistressParticlesNoSearch | *Probability without search* of being in this scenario and in distress |
| jointDistressParticlesWithSearch | *Probability with search* of being in this scenario and in distress |
| jointLandedParticlesNoSearch | *Probability without search* of being landed (and in distress) and in this scenario |
| jointLandedParticlesWithSearch | *Probability with search* of being landed (and in distress) and in this scenario |
| landedCount | Number of particles in this scenario that are landed (and in distress) |
| Conditionals (conditioned on scenario) | |
| landedParticlesNoSearch | *Probability without search* of being in this scenario, given that it is landed |
| landedParticlesWithSearch | *Probability with search* of being in this scenario, given that it is landed |

We propose the elimination of the entire SCENARIO tag. To my knowledge, it is not being used.

##### POS Tags

Also within the “TIME” tag, there is a “POS” subtag for each search object type. A subtag of the “POS” subtag gives geographic information about the particles that are (at this time) of that search object type. In the “POS” tag itself, there are the following entries:

|  |  |
| --- | --- |
| conditionalPOS | Probability of finding the object, given that it is this search object type |
| initialProbability | *Probability without search* of being this search object type |
| jointPOS | Probability of finding the object *and* it is this search object type. |
| numberInDistressAndLanded | Number of particles of this search object type that are in distress and landed |
| numberInDistressAndNotLanded | Number of particles of this search object type that are in distress and not landed |
| remainingProbability | Probability without search of being this search object type, minus the probability of success, given that the target is this search object type. |
| remainingTargetProbability | Same as remainingProbability. The two attributes are here for backward compatibility reasons. |

We propose the elimination of the remainingTargetProbability attribute. To my knowledge, it is not being used.

## Cleanup of the Statistics File

SAROPS produces a file called “Stats.” We propose the elimination of the this entire file. To my knowledge, it is not being used.