



CADSWES University of Colorado

Center for Advanced Decision Support for Water and Environmental Systems

MRM and Post-Processing

RiverWare User Group Meeting
February 10-11th, 2010

Bill Oakley

MRM Introduction

- Reclamation uses Multiple Run Management Concurrent Mode with the CRSS model for policy and planning purposes
- Availability of Paleo-Data has greatly increased the required number of traces (~1200)
(A trace is a single simulation within a multiple run)
- Tasked with helping Reclamation run the required number of traces

MRM Challenges

➤ Significant challenges

- Memory – Small per-trace memory growth limited a multiple run to ~125 traces before an out-of-memory error
- Performance – Estimated 20 hours for 1200 traces

MRM Memory Growth

- Two common causes of memory growth
 - Memory leaks
 - ❖ Caused by allocating memory and not deleting it
 - ❖ Rational Purify memory analysis tool eliminated memory leaks as a cause of the memory growth
 - Memory fragmentation
 - ❖ Caused by excessive memory allocations and deletions
 - ❖ Several strategies to reduce memory fragmentation, one of which is, not surprisingly...


MRM Memory Fragmentation



Reducing memory
allocations and
deletions

Captain Obvious

MRM Memory Fragmentation

- RiverWare maintains data structures whose size is determined by workspace attributes (number of objects, number of timesteps, etc)
- Typically the data structures are deleted and reallocated before a simulation
- During a multiple run the workspace attributes are static – data structures can be reused 

MRM Memory Fragmentation

- Modified the code to reuse data structures during a multiple run
 - 932 traces before an out-of-memory error
 - 65 seconds per trace → 21.6 hours for 1200 traces
- At this point we had picked the low hanging fruit in the memory fragmentation orchard, and decided it would not be cost beneficial to pursue memory fragmentation further

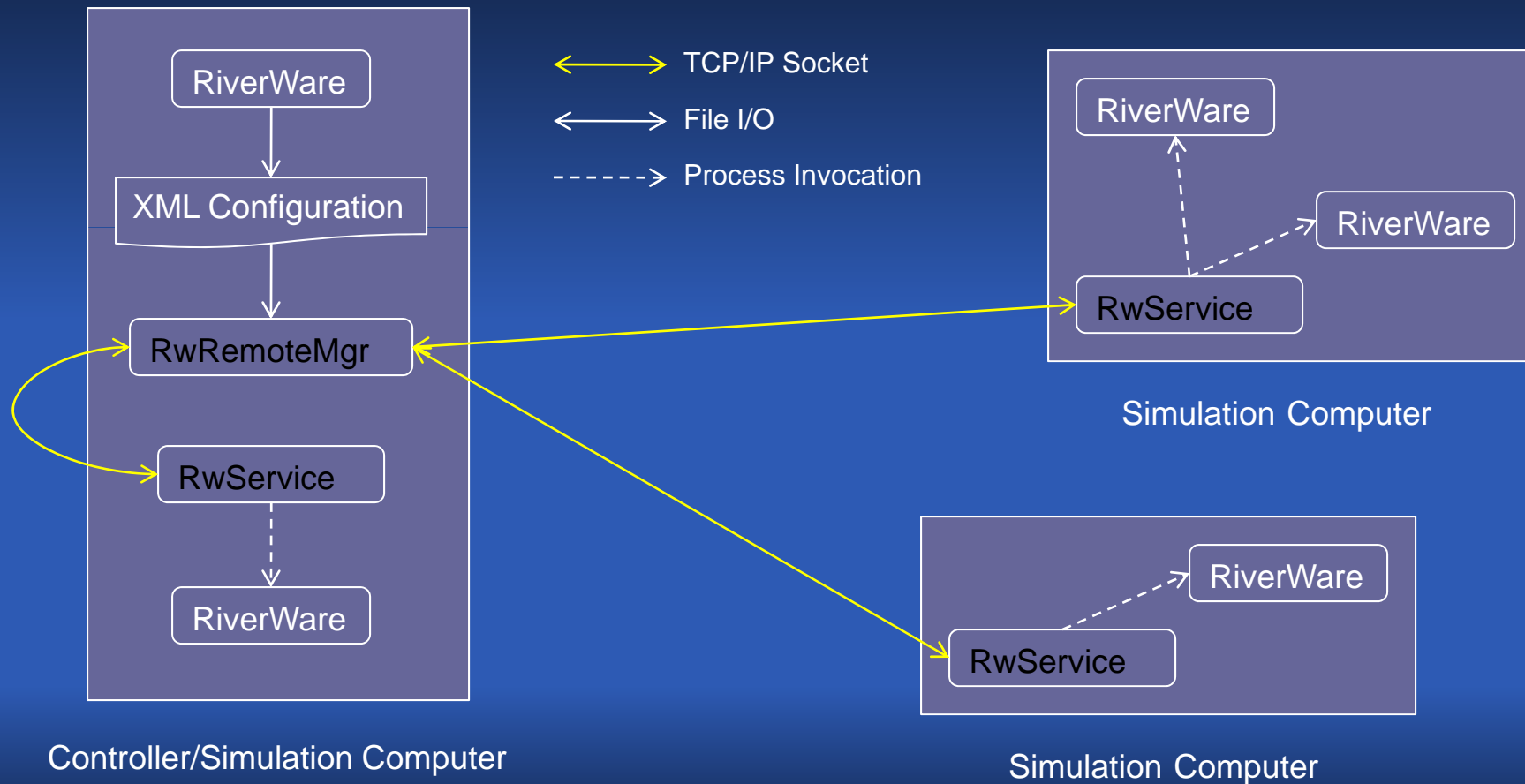
MRM Performance

- Two pronged approach
 - Analyze RiverWare code with Rational Quantify performance analysis tool
 - Analyze ruleset (bring it up to date with latest RPL enhancements)
- RiverWare code changes and ruleset changes
- Significant improvement (23%)
 - 50 seconds per trace → 16.6 hours for 1200 traces
- Time for “Plan B”

MRM “Plan B”

- In a concurrent multiple run each trace is independent of all other traces
- Conceptually the traces can be distributed across multiple computers, each performing a subset of the traces
- Implemented Distributed Concurrent Multiple Runs

Distributed MRM Architecture



Distributed MRM Configuration

escription Output Run Parameters Policy Input Distributed Runs

Login As: User Password Not Secure [Learn More...](#)

Working Directory:

Save Distributed Configuration As:

Simulations

Distribute Evenly Port: Number of Traces: 1200 of 1200

	Host	First Trace	Last Trace	Num Traces
	training0	1	110	110
	training1	111	219	109

Environment Variables

	Variable	Value
	CRSS_DIR	C:/CRSS

Remote Manager

The screenshot displays the RiverWare Distributed MRM interface for a training session. The window title is "RiverWare Distributed MRM - Baseline d12 1200t 51y training". The interface is organized into two rows, one for host "training13" and one for host "training14". Each row contains three panels: "Process Status", "Multiple Run Status (Traces 1001 - 1100)", and "Single Run Status".

Host: training13

- Process Status:** Host: training13, State: Running. Includes play, stop, and log icons.
- Multiple Run Status (Traces 1001 - 1100):** Progress bar at 14%. Execution State: Running. Current Run: 15 of 100.
- Single Run Status:** Progress bar at 51%. Execution State: Running. Current Timestep: April, 2036.

Host: training14

- Process Status:** Host: training14, State: Running. Includes play, stop, and log icons.
- Multiple Run Status (Traces 1101 - 1200):** Progress bar at 14%. Execution State: Running. Current Run: 15 of 100.
- Single Run Status:** Progress bar at 25%. Execution State: Running. Current Timestep: January, 2023.

Estimated Time Remaining: 02:43:54

Start All Stop All

Remote Manager

The screenshot displays the RiverWare Distributed MRM interface. The main window is titled "RiverWare Distributed MRM - Baseline d12 1200t 51y training". It features two panels for host status:

- Host: training13**
 - State: Running
 - Multiple Run Status (Traces 1001 - 1100): 7% progress, Execution State: Running, Current Run: 8 of 100.
 - Single Run Status: 100% progress, Execution State: Running, Current Timestep: December, 2060.
- Host: training14**
 - State: Running
 - Multiple Run Status (Traces 1101 - 1200): 7% progress, Execution State: Running, Current Run: 8 of 100.
 - Single Run Status: 100% progress, Execution State: Running, Current Timestep: December, 2060.

An estimated time remaining of 02:53:49 is shown at the bottom left. A tooltip "View RiverWare Diagnostic Output" points to a document icon in the training13 host status panel.

The "RwRemoteMgr" window in the foreground shows the following diagnostic output:

```
_REQINFO_: "----- Rulebased Simulation RUN FINISHED (MRM run 100 of 100) -----"  
_REQINFO_: ""CRSS.2010.BiNational.mdl at 20:14 February 5, 2010 (100 seconds)""  
_REQINFO_: "-----"  
_REQINFO_: "----- MRM RUN FINISHED -----"  
_REQINFO_: ""CRSS.2010.BiNational.mdl at 20:14 February 5, 2010 (11329 seconds)""  
_REQINFO_: "-----"
```

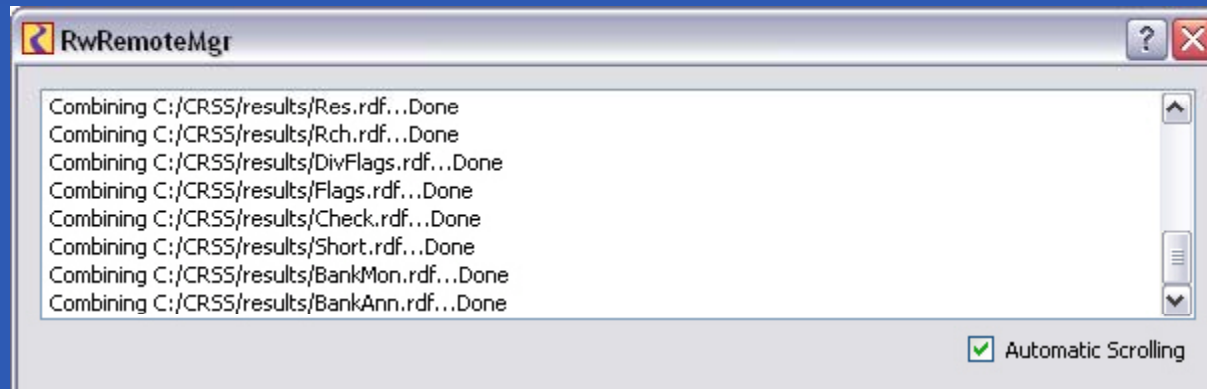
The "Automatic Scrolling" checkbox is checked at the bottom right of the diagnostic window.

Remote Manager

Post Processing: Combining RDF file Div.rdf



26%



Distributed MRM Performance

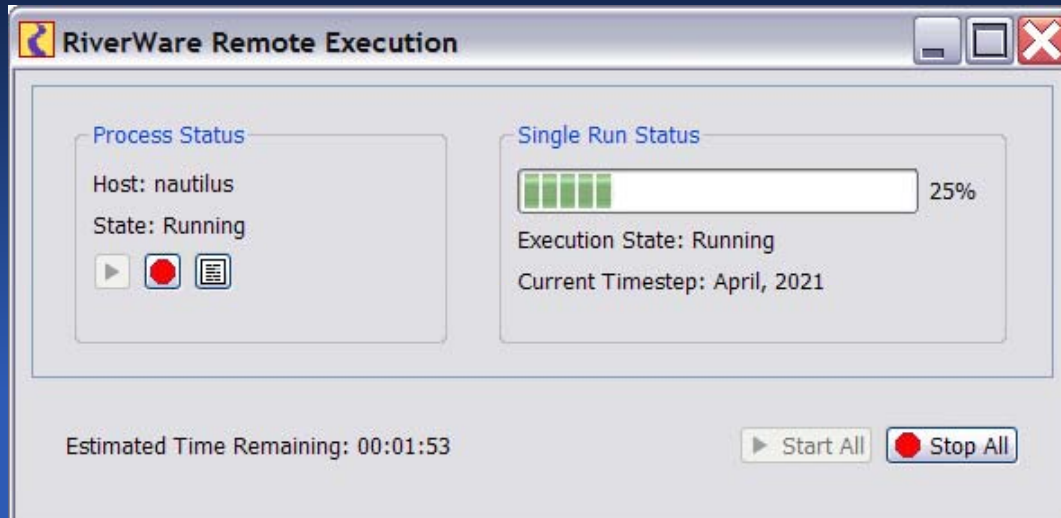
- 1200 traces distributed across 11 computers
 - 3 hours, 37 minutes
- Can always lower run time by applying more, or faster, computers

Remote RiverWare Execution

➤ Controlled by XML configuration file

```
<document>
  <RW>
    <host addr="training1" port="27285"/>
    <app>c:/Program Files/CADSWES/RiverWare 5.2/riverware.exe</app>
    <script name="c:/models/crssShortageDEIS.rcl">
      <openws>c:/models/crssShortageDEIS.mdl.gz</openws>
      <loadrules>c:/models/crssShortageDEIS.rls</loadrules>
      <start/>
      <close/>
    </script>
    <output>c:/models/crssShortageDEIS.log</output>
    <envlist>
      <env>RIVERWARE_HOME_52=c:\\Program Files\\CADSWES\\RiverWare 5.2</env>
    </envlist>
  </RW>
</document>
```


Remote RiverWare Execution



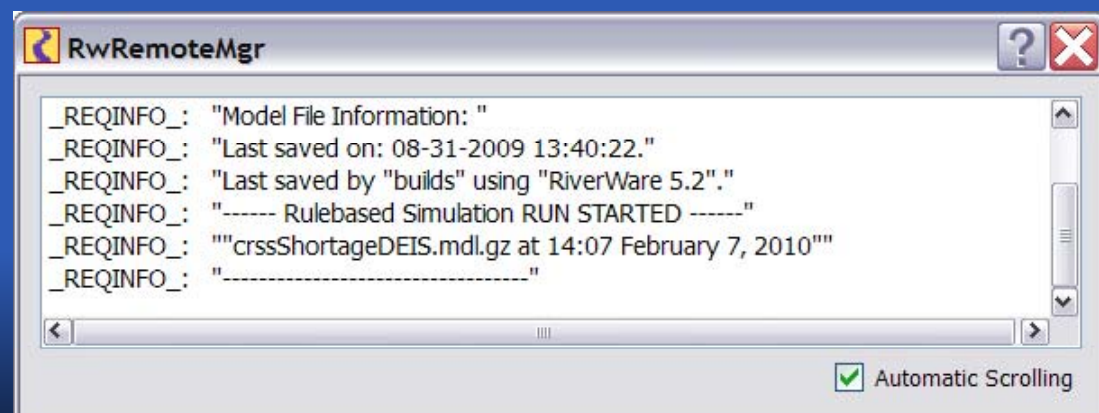
RiverWare Remote Execution

Process Status
Host: nautilus
State: Running
▶ ● 📄

Single Run Status
25%
Execution State: Running
Current Timestep: April, 2021

Estimated Time Remaining: 00:01:53

▶ Start All ● Stop All



RwRemoteMgr

```
_REQINFO_: "Model File Information: "  
_REQINFO_: "Last saved on: 08-31-2009 13:40:22."  
_REQINFO_: "Last saved by "builds" using "RiverWare 5.2"."  
_REQINFO_: "----- Rulebased Simulation RUN STARTED -----"  
_REQINFO_: ""crssShortageDEIS.mdl.gz at 14:07 February 7, 2010""  
_REQINFO_: "-----"
```

Automatic Scrolling

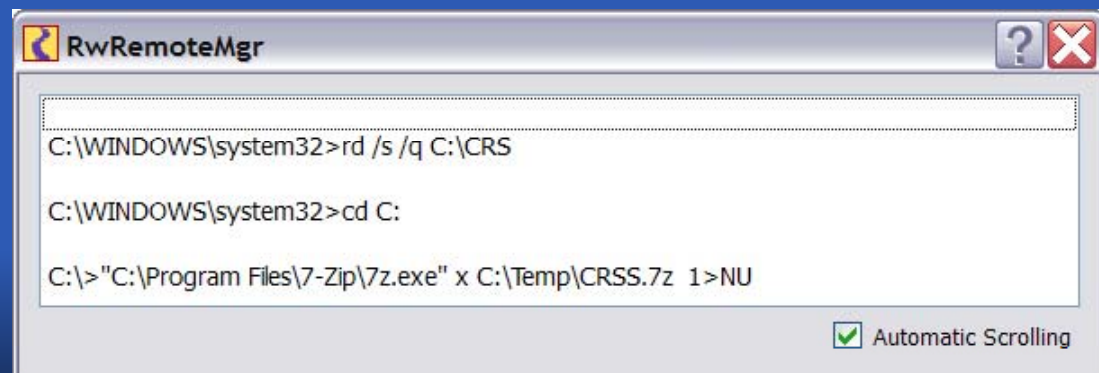
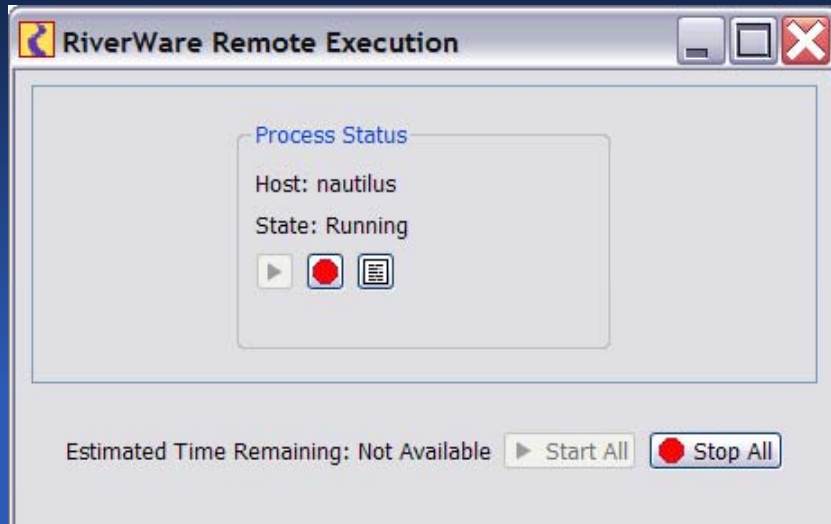
Remote Application Execution

- Also controlled by XML configuration file

```
rd /s /q C:\CRSS
copy \\animas\models\CRSS.7z C:\Temp
cd C:\
"C:\Program Files\7-Zip\7z.exe" x C:\Temp\CRSS.7z > NUL
del C:\Temp\CRSS.7z
```

```
<document>
  <GenApp>
    <host addr="training1" port="27285"/>
    <app>\\animas\models\copyCRSS.bat</app>
  </GenApp>
</document>
```

Remote Application Execution



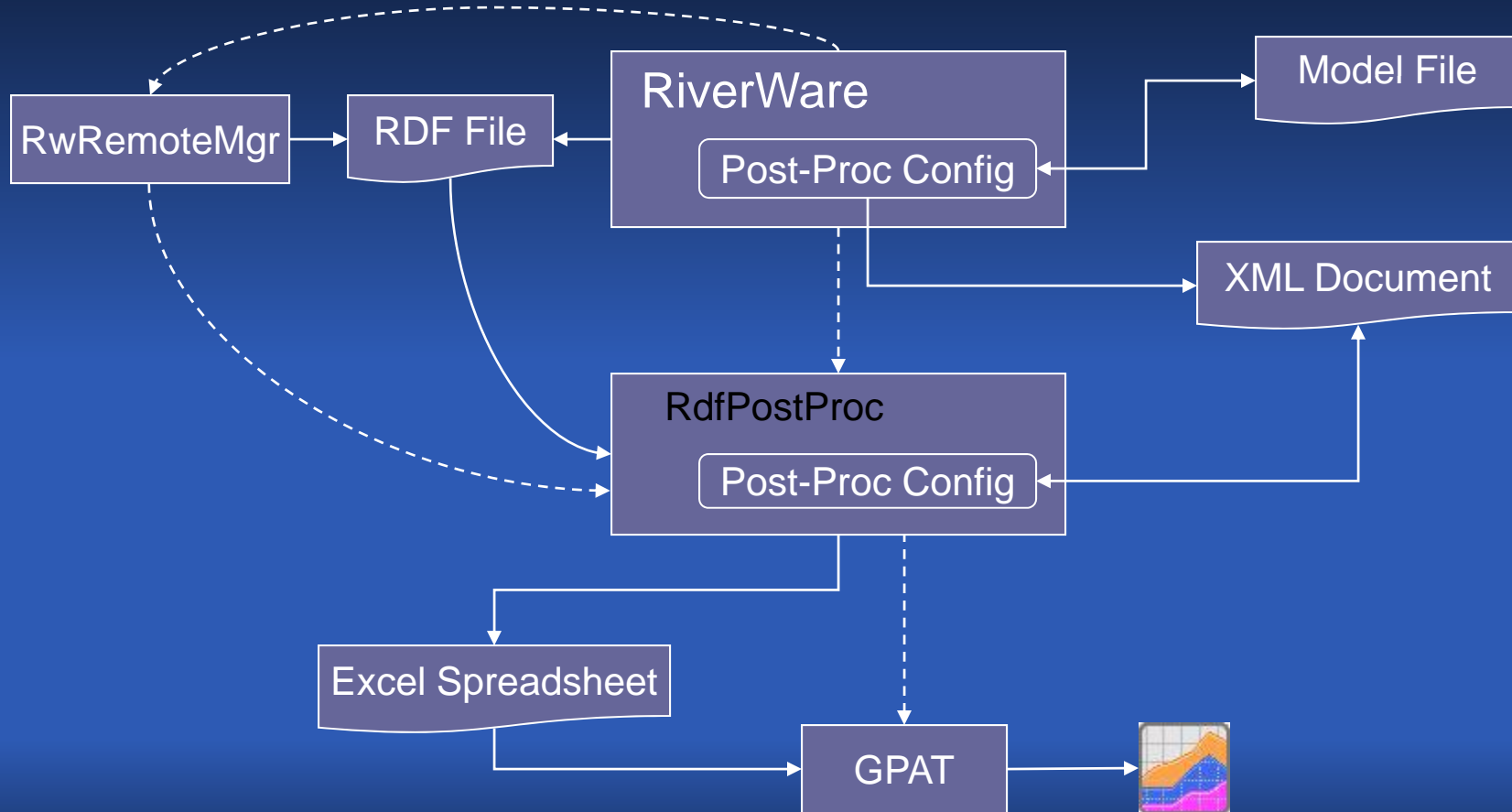
Post-Processing

- Multiple run output is RDF files
- Post-processing currently includes
 - Transforming RDF files (e.g., transforming a monthly RDF file to an annual RDF file)
 - Generating non-RDF output from RDF files (e.g., Excel spreadsheets)
 - Generating plots from Excel spreadsheets (using GPAT, an Excel add-in written in VBA)

Post-Processing

- Tasked with helping Reclamation automate their post-processing
 - Click the “Start” button, come back in 3 hours and 37 minutes and have RDF files, Excel spreadsheets and plots (as graphics files) ready to go
- Solution must be easily extended to provide other transform and output capabilities
 - Plugins – define transform and output plugin interfaces; DLLs which implement the interface can be dynamically added to provide additional post-processing capabilities

Post-Processing



Post-Processing

- Requirements and high-level design phase
 - RiverWare – Configure post-processing
 - RdfPostProc – Replace ExcelWriter and YearlyAgg, configure post-processing, generate Excel spreadsheets, invoke GPAT in batch mode to generate plots (as graphics files)
 - GPAT – Stored plots, batch mode
 - RDF file format changes?



CADSWES University of Colorado

Center for Advanced Decision Support for Water and Environmental Systems

Graphical Policy Analysis Tool (GPAT)

RiverWare User Group Meeting
February 10-11th, 2010

Tim Magee & Neil Wilson

Why GPAT?

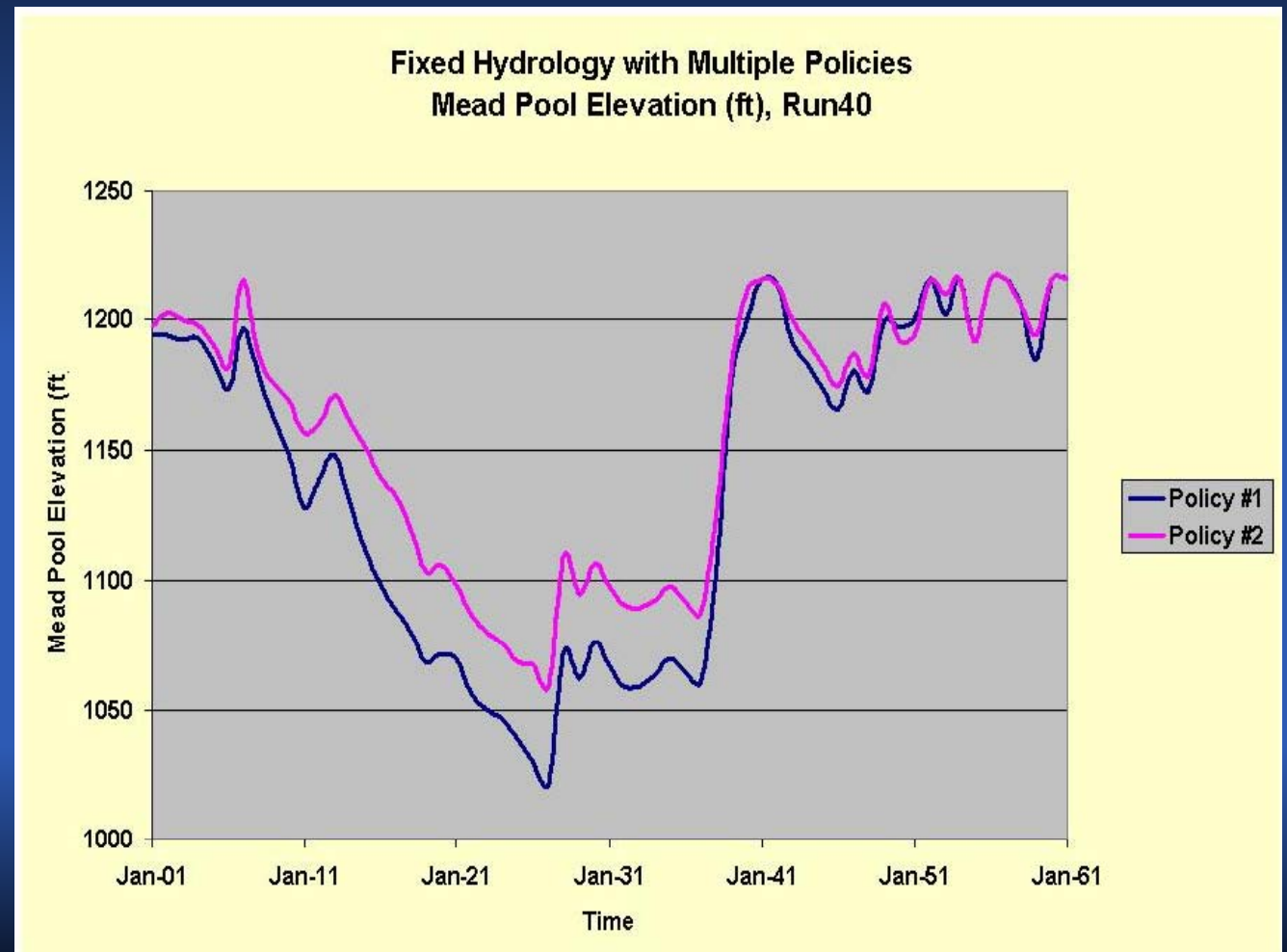
- RiverWare Output Large, esp.
 - Alternative Hydrologic Scenarios / Histories: MRM
 - Alternative policies
- Information from Data
 - Statistics
 - Graphs
 - Dynamic generation
- Presentations and Documents
 - e.g. Stakeholders meetings and EIS

What is GPAT?

- Excel Add-in written in VBA
 - Download from the RiverWare web site
- Inputs are Excel Workbooks
 - Dimensions are Slots, Time, Runs, and Policies
 - Variety of orientations
- Outputs are Charts and Worksheets in a Workbook
 - Typical Chart: several statistical time series
 - Present directly or paste into documents

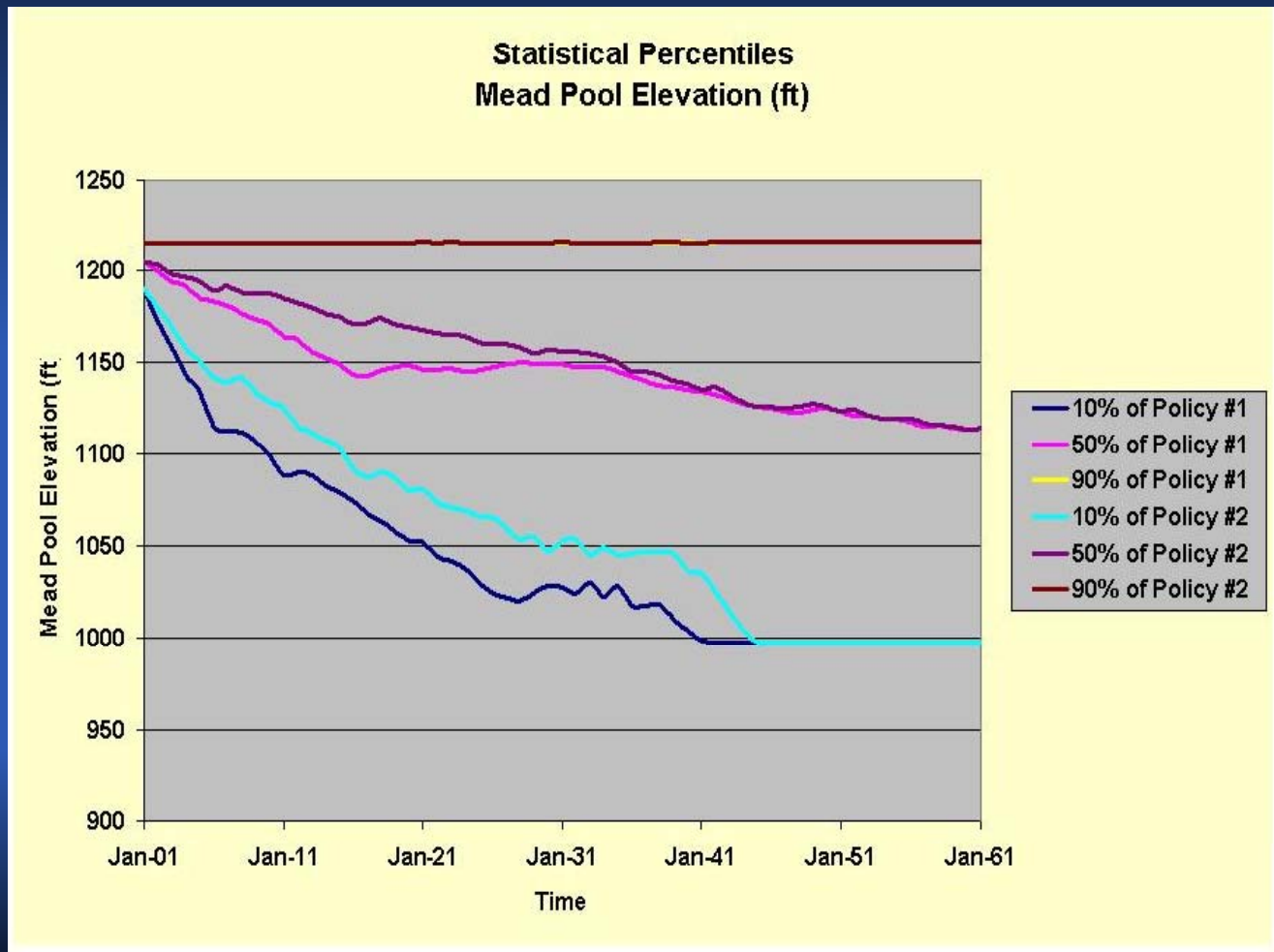
Compare Traces

- Compare traces under different policies



Simple Statistics

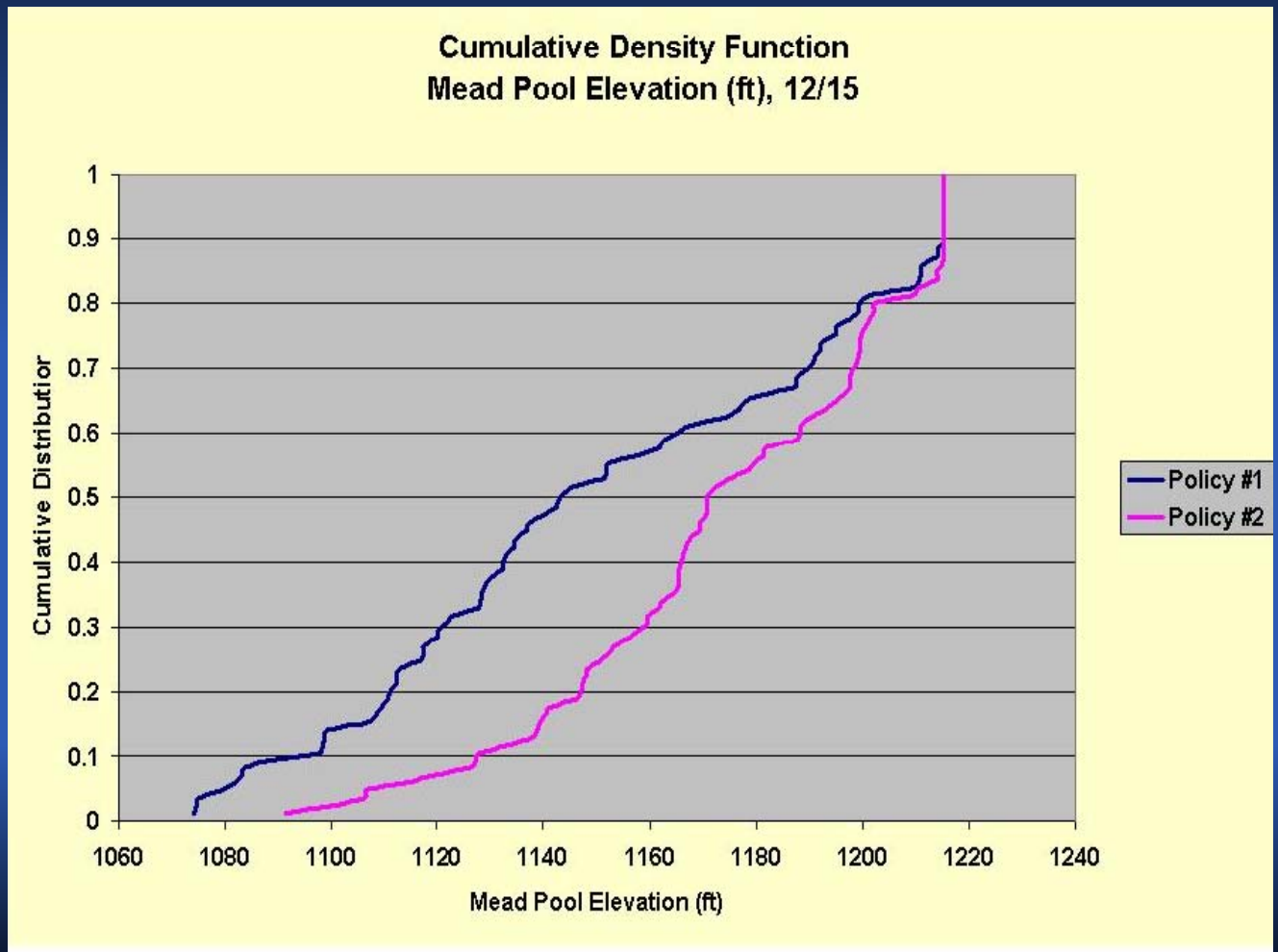
- Minimum
- Maximum
- Mean
- Std. Dev.
- Median
- Percentiles
 - Alternative Methods



Distributions

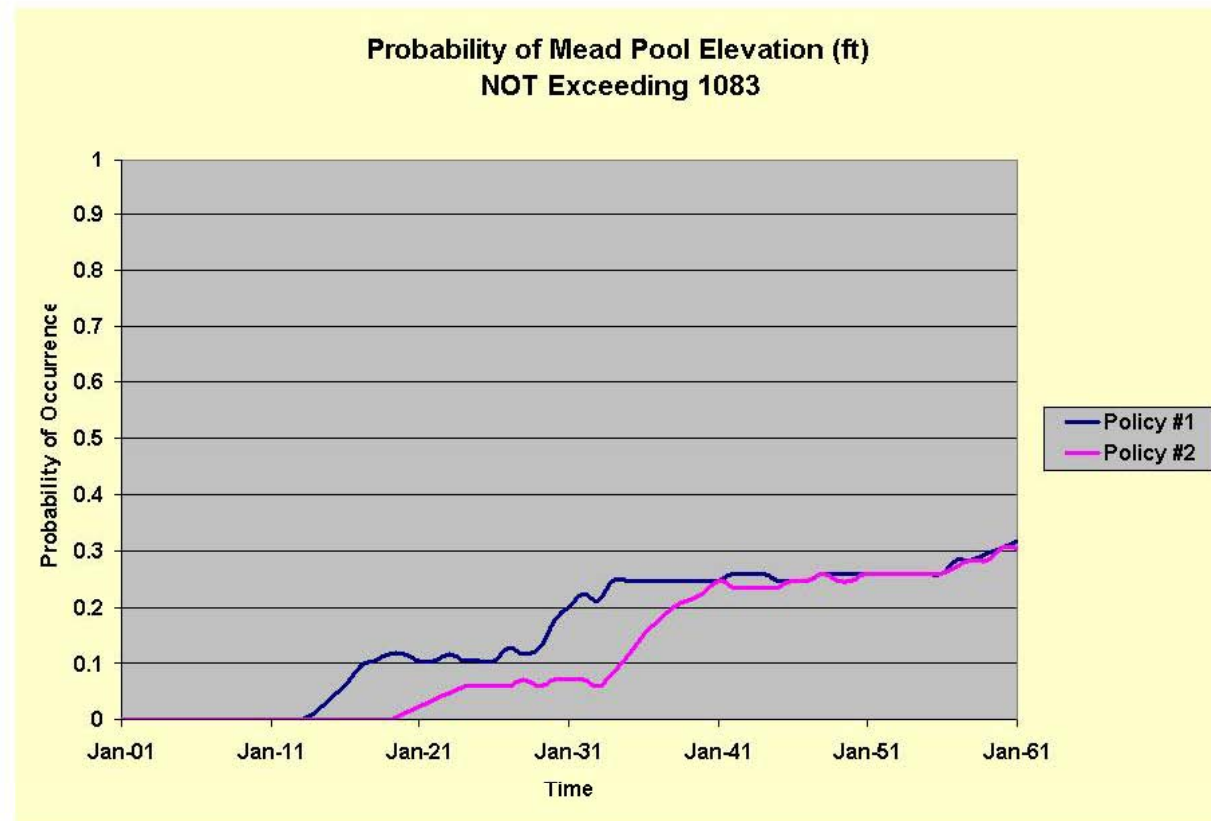
Histogram CDF

- One date
- All time
- Annual peak
- Maximum for a given duration



Probability of Events

- Binary events
- Exceedance
- Occurrence in a range
- Compound Events



Upcoming Work

- Connection to RiverWare - Bill
- Box plots?
 - Manually feasible!
 - Programmatically feasible?
Investigation
 - Implementation if feasible

