WRF #4742: Probability Management for Water Utilities

Research Objectives

1. Principles of Probability Management (PM)
2. Illustrate the use of PM tools in depicting and communicating uncertainty
3. SIPmath™ v2 standards

Funders

- Water Research Foundation
- Metropolitan Water District of Southern California
- Irvine Ranch Water District
- Tacoma Water
- Inland Empire Utility Agency
- Eastern Municipal Water District

Research Team

- ***Dr. Thomas W. Chesnutt, PStat® , CAP®*** (Principal Investigator) of A&N Technical Services,
- ***Dr. Michael Hollis, PStat®*** of the Metropolitan Water District of Southern California,
- ***Shayne Kavanagh*** of the Government Finance Officers Association,
- ***David L. Mitchell*** of M. Cubed,
- ***Dr. David M. Pekelney*** of A&N Technical Services,
- ***Dr. Jean-Daniel Rinaudo*** (Outside Expert), of the French Geological Survey (BRGM), and
- **Marc Thibault** the lead author of SIPmath™ v2 standards for probabilitymanagment.org.

Room for improvement

Are studies reproductible?

The New York Times

TheUpshot

THE NEW HEALTH CARE

Science Needs a Solution for the Temptation of Positive Results

By Aaron E. Carroll

May 29, 2017

153

A few years back, scientists at the biotechnology company Amgen set out to replicate 53 landmark studies that argued for new approaches to treat cancers using both existing and new molecules. They were able to replicate the findings of the original research only 11 percent of the time.

Science has a reproducibility problem. And the ramifications are widespread.

Reliance on Averages

- Good decisions need to account for:
  - the entire shape of uncertainty (***distributions***) ; and
  - the consequences of risky outcomes (***loss function***).
- Averages obscure both, leading to mis-informed decisions.

WRF 4558 Uncertainty in Long-Term Water Demand Forecasts - Survey Responses

- Majority are 'cognizant of long term demand forecast uncertainties' and have taken some steps to address
- Larger systems (pop. served) means more attention devoted to long-term water demand forecasting
- Key future uncertainties: future growth in customers, future climate, the condition of the economy, and water efficiency
- ~50% of the utilities that address uncertainty do not address all variables
- There are deficiencies in available data and forecasts of available data
- Water supply & distribution system planning
  - bias for demand forecast scenarios reflect middle-to-high range of predictions,
  - under-predicting future demands outweigh the risks of over-predicting

ProbabilityManagement.org

- Non-profit
- Aim
  - Improve use of uncertainty in decision making
  - “communication, calculation, and credible estimates.”
- Data array aka Stochastic Information Packet
  - Define it
  - Way to contain it and use across software platforms
  - Meta data
  - Standard
- Free tools, purchase tools

Principles of Probability Management

(See probabilitymanagement.org)

- **Communicating Distributions as Data** –When estimating uncertain quantities, a "typical" or "average" value is used. Collapsing uncertainty to a scalar destroys information on variability.
- **Information** (measurement) can reduce total uncertainty and add credibility.
- **Interactive simulation** is a useful tactic for gaining simulated experience with how decisions affect risky outcomes. Visualization is an effective medium for communicating risk to decision-makers.
- **Coherence**—Distributions of causal forces are often related/dependent. The SIPmath Standard provides the **glue** logic for related uncertainties.

Probability Management's *SIPmath™ 2.0 Standard* defines:

SIP – Stochastic Information Packets

- *Uncertainties are communicated as data arrays*
- *Thus random draws from uncertain possibilities are stored as a column of realizations.*

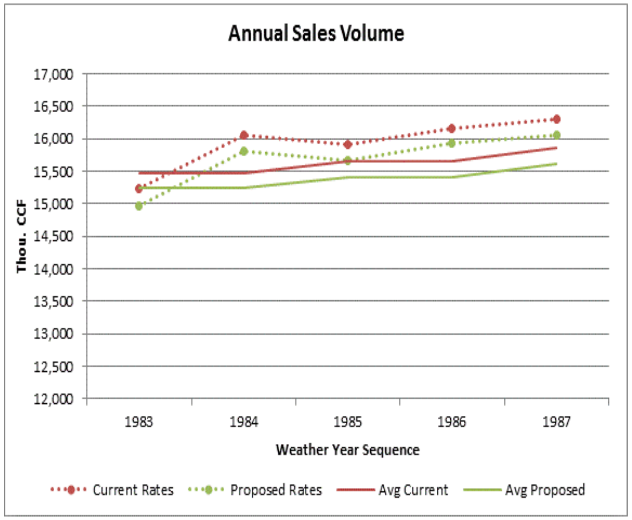
SLURP – Stochastic Library Unit Relations Preserved

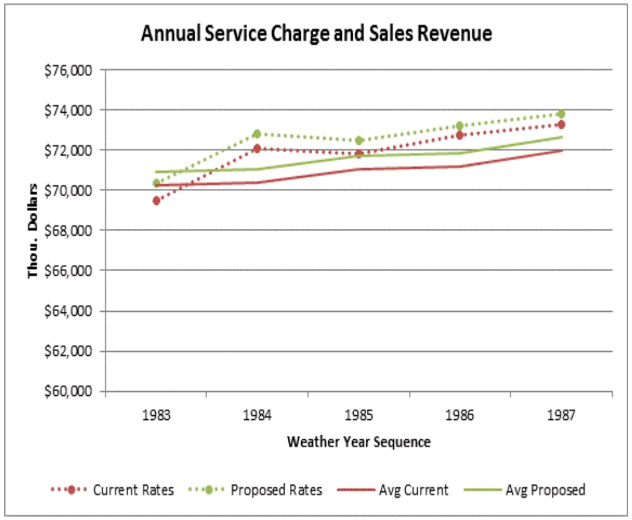
- *A **coherent** set of SIPs that preserve statistical relationships between uncertainties*

SIPMath™ is Actionable

Sips can be used directly in calculations of uncertainty.
Cells in Excel can refer to SIPs instead of a single number.
No macros or add-ins need remain in the spreadsheet.

Volume and Revenue Volatility demonstrated visually





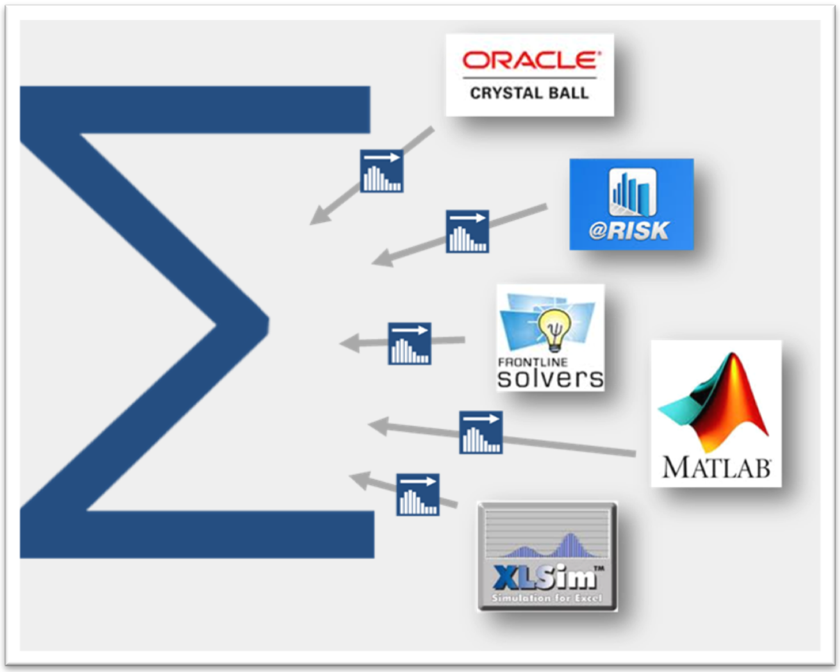
SIPmath compliant open-source model at http://www.financingsustainablewater.org/tools/awe-sales-forecasting-and-rate-model

SIPMath™ is Additive

Uncertainties can be summed, enabling enterprise risk management

SIPmath

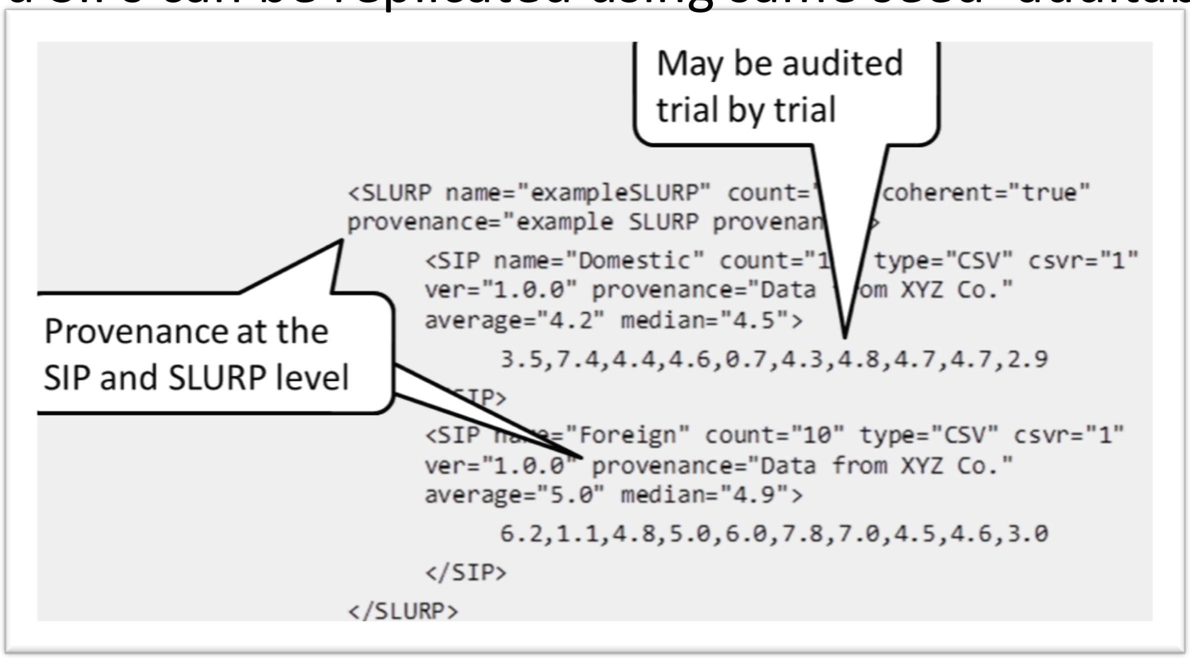
- Platform independent
- Excel add in
    - free at probability management
    - any of these commercial simulation software.



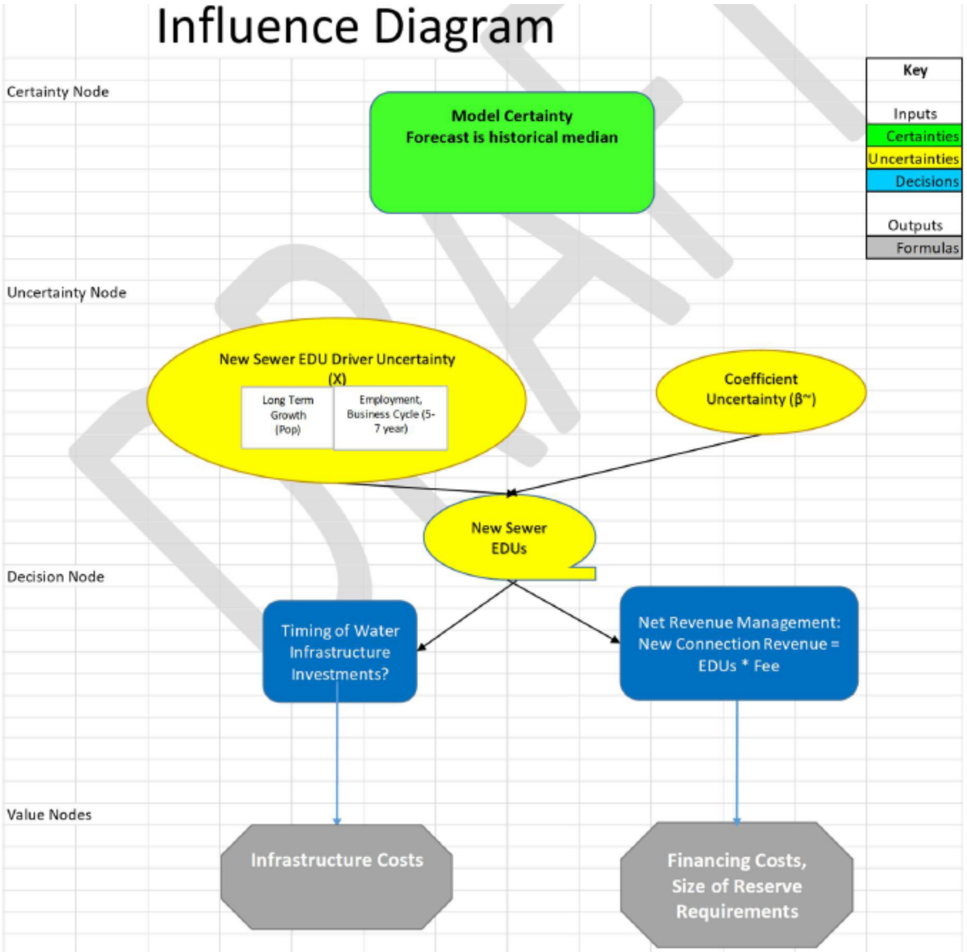
SIPMath™ is Auditable

The SIPMath™ standard requires provenance.

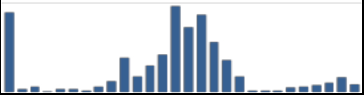
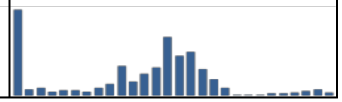
Saved SIPs can be replicated using same seed=auditability.



New Connection Revenue



Predicting New Connection Revenue

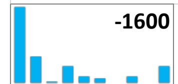
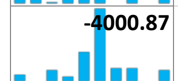
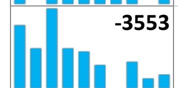
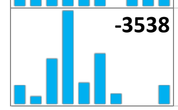
NewSewerEDUs		%GrowthSewerEDUs
3,116	median	1.71%
5,586	mean	3.14%
4,985	std dev	3.33%
		

Four Models of the Annual Change in EDUs: Advantages and Disadvantages

Model Key	Model Description	Advantages	Disadvantages
Model 0	Pick the historically most likely (median ~= 3,260 Delta_EDU)	Easiest to implement Easy to explain Overprediction can never be large	Does not predict booms or busts
Model 1	Pick year before Delta_EDU	Easy to implement/explain Captures some boom/bust Small errors (<2K) 60% of time	Most variable prediction error, on average Large errors (>2K EDU) one third of the time
Model 2	Lagged Percentage Growth of Employment	Relatively simple model Captures some boom/bust	Requires employment data and implementation
Model 3	year before Delta_EDU + Year before change in Emp growth + Difference of Emp growth from trend	Best predictor, on average	Not easy to implement or explain large errors still possible

25%

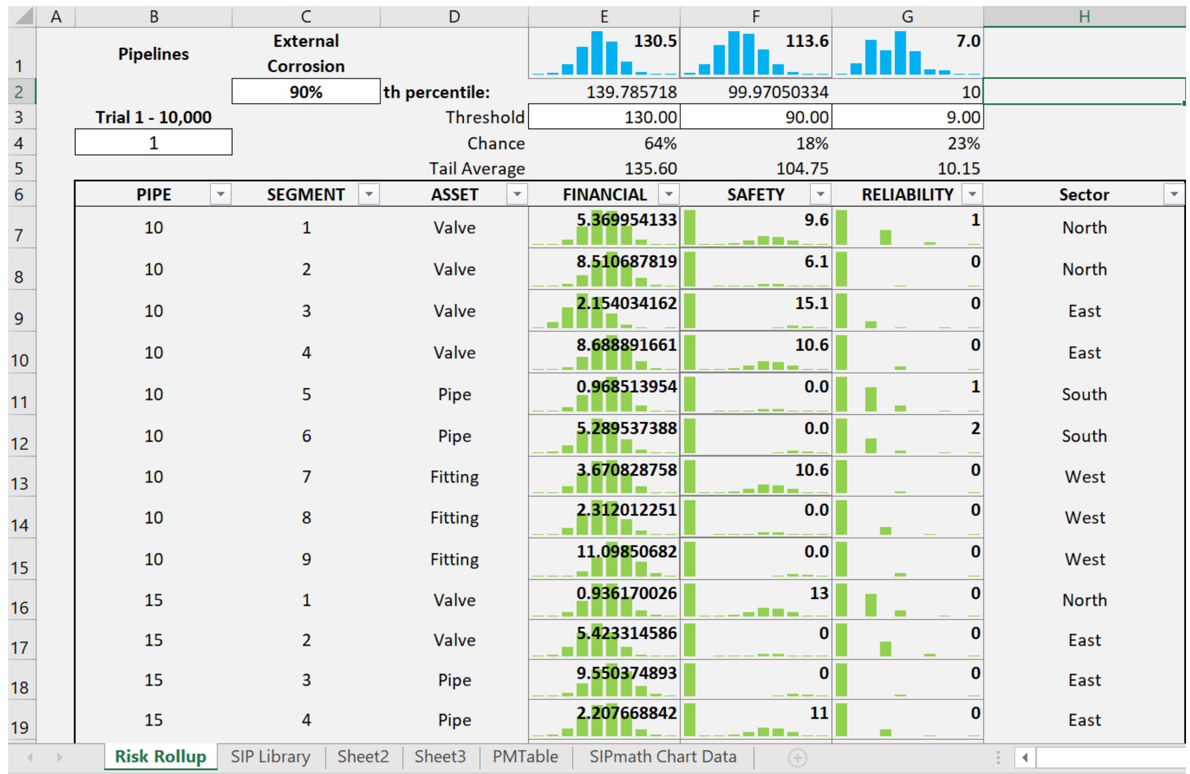
60%

		Variability	Median
ModelError=Actual-Predicted (Y-Y_hat)		Std.Dev.	50%
 -1600	d_Model0Error	5,060	-
 -4000.87	d_Model1Error	5,252	181
 -3553	d_Model2Error	5,056	(471)
 -3538	d_Model3Error	4,660	(691)

- # new connections each year is very uncertain on a one year time horizon.
- Explore the models

SIPmath Example for Diverse Risk Aggregation in a Gas Utility Network

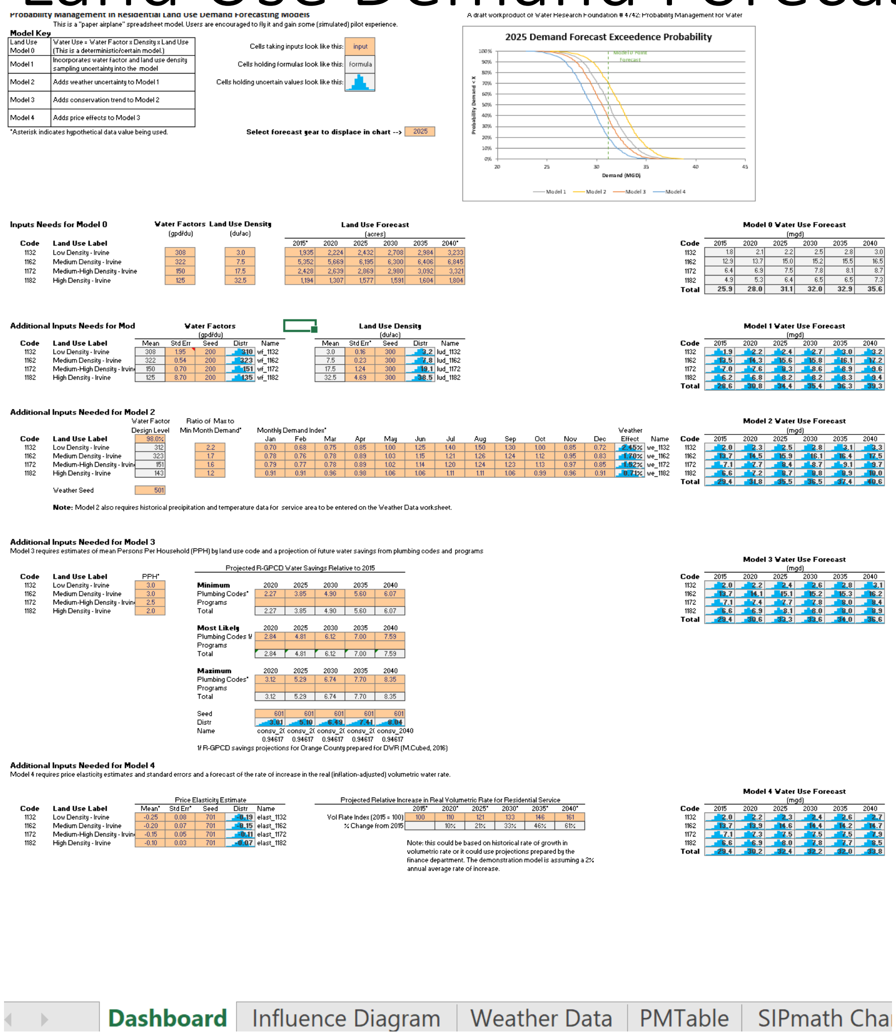
- Aggregating operational risk across physical assets
- Risk Mitigation investment in a highly regulated environment require stakeholder understanding of Cost – Risk Tradeoffs across different types of risk: financial, safety, and reliability



	A	B	C	D	E	F	G	H
1		Pipelines	External Corrosion		130.5	113.6	7.0	
2			90%	th percentile:	139.785718	99.97050334	10	
3		Trial 1 - 10,000		Threshold	130.00	90.00	9.00	
4		1		Chance	64%	18%	23%	
5				Tail Average	135.60	104.75	10.15	
6		PIPE	SEGMENT	ASSET	FINANCIAL	SAFETY	RELIABILITY	Sector
7		10	1	Valve	5.369954133	9.6	1	North
8		10	2	Valve	8.510687819	6.1	0	North
9		10	3	Valve	2.154034162	15.1	0	East
10		10	4	Valve	8.688891661	10.6	0	East
11		10	5	Pipe	0.968513954	0.0	1	South
12		10	6	Pipe	5.289537388	0.0	2	South
13		10	7	Fitting	3.670828758	10.6	0	West
14		10	8	Fitting	2.312012251	0.0	0	West
15		10	9	Fitting	11.09850682	0.0	0	West
16		15	1	Valve	0.936170026	13	0	North
17		15	2	Valve	5.423314586	0	0	East
18		15	3	Pipe	9.550374893	0	0	East
19		15	4	Pipe	2.207668842	11	0	East

Reference: [Rolling up operational risk at PG&E]()
by Jordan Alen, Christine Cowsert Chapman, Melissa Kirmse, Farshad Miraftab, & Sam L. Savage. OR/MS Today, December 2016, Volume 43, Number 6.
The SIPmath paper airplane "Asset Level Model.xlsx" demonstrates how different types of risks—financial, safety, and reliability—can be combined to show cost tradeoffs.

Land Use Demand Forecasting



Probability Management in Residential Land Use Demand Forecasting Models

This is a "paper airplane" spreadsheet model. Users are encouraged to fly it and gain some (simulated) pilot experience.

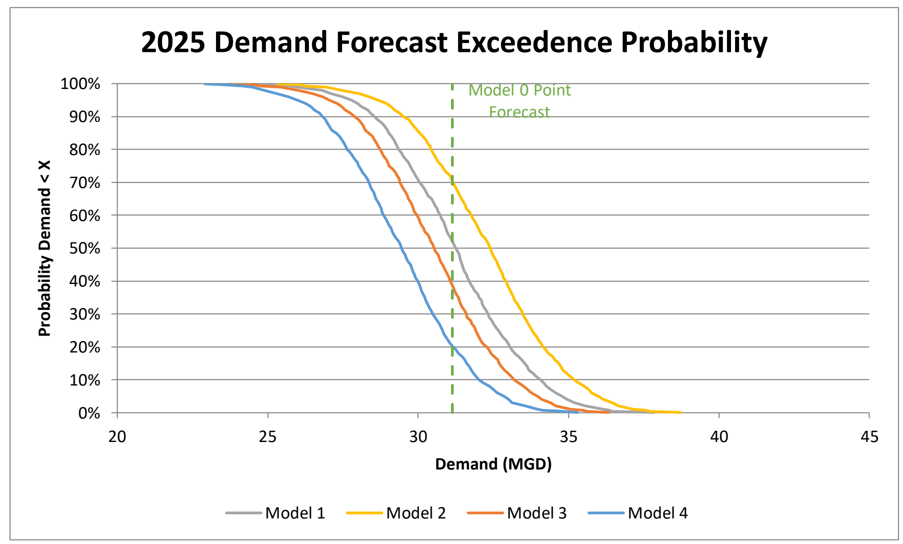
Model Key

Land Use Model 0	Water Use = Water Factor x Density x Land Use (This is a deterministic/certain model.)
Model 1	Incorporates water factor and land use density sampling uncertainty into the model
Model 2	Adds weather uncertainty to Model 1
Model 3	Adds conservation trend to Model 2
Model 4	Adds price effects to Model 3

Cells taking inputs look like this: input

Cells holding formulas look like this: formula

Cells holding uncertain values look like this:



Conclusion

- By reducing the cost of creating, validating, and communicating probabilistic analysis... Probability Management gives permission to face uncertainty and inform decisions.
- Ignoring uncertainty does not make it go away. It only leads to ulcers and sleepless nights.

Resources:

- Probabilitymanagement.org
- Open source risk based forecast book, http://www.gfoa.org/forecastbook
- AWE, www.financingsustainablewater.org