What Remains to be Done with Demand Response? A National Forum from the FERC National Action Plan on Demand Response Tries to Give an Answer

Measurement and Verification Working Group

Miriam Goldberg
Senior Vice President
DNV KEMA Energy and Sustainability

National Association of Regulatory Utility Commissioners, ER&E Committee Meeting, July 24, 2012 Portland, OR

M&V Working Group Mission

- Scope
  - Review work to date to establish DR M&V protocols and baseline calculation methods
  - Identify
    - Accepted methods and practices for DR performance measurement
    - Areas still at issue
    - Gaps related to protocols and practices for specific types of DR programs, emerging technologies, or markets
  - Provide a path forward for industry and stakeholders towards analytically valid, widely accepted DR M&V protocols or best practices.
- Emphasis: calculation of DR performance
  - Wholesale market settlement
    - Between ISO and market participant or ISO program participant
  - Retail settlement
    - DR aggregator or utility program operator and retail customer
  - Program impact evaluation
    - for retail or ISO program
- Guidance, not standards

M&V Working Group Members

DNV KEMA	DNV KEMA Energy and Sustainability	Mimi Goldberg, Chair
ICF	*ICF International*	*Kara Nave, Logistics*
LBNL	*Lawrence Berkeley National Laboratory*	*Liz Stuart, Coordinator*
BG&E	BGE, an Exelon Company	Cheryl Hindes
CEC	California Energy Commission	David Hungerford
ERE	Energy Resource Economics	Theresa Flaim
EnerNOC	EnerNOC Solutions Consulting	Craig Williamson
FSC	Freeman, Sullivan & Co.	Steve George
ISO NE	Independent System Operator New England	Henry Yoshimura
NEEP	Northeast Energy Efficiency Partnerships	Julie Michals, Elizabeth Titus
NPCC	Northwest Power and Conservation Council	Ken Corum
NY ISO	New York Independent System Operator	Donna Pratt
OPA	Ontario Power Authority	Nik Schruder
PG&E	Pacific Gas and Electric Company	Mike Alexander
SCE	Southern California Edison	Mark Martinez

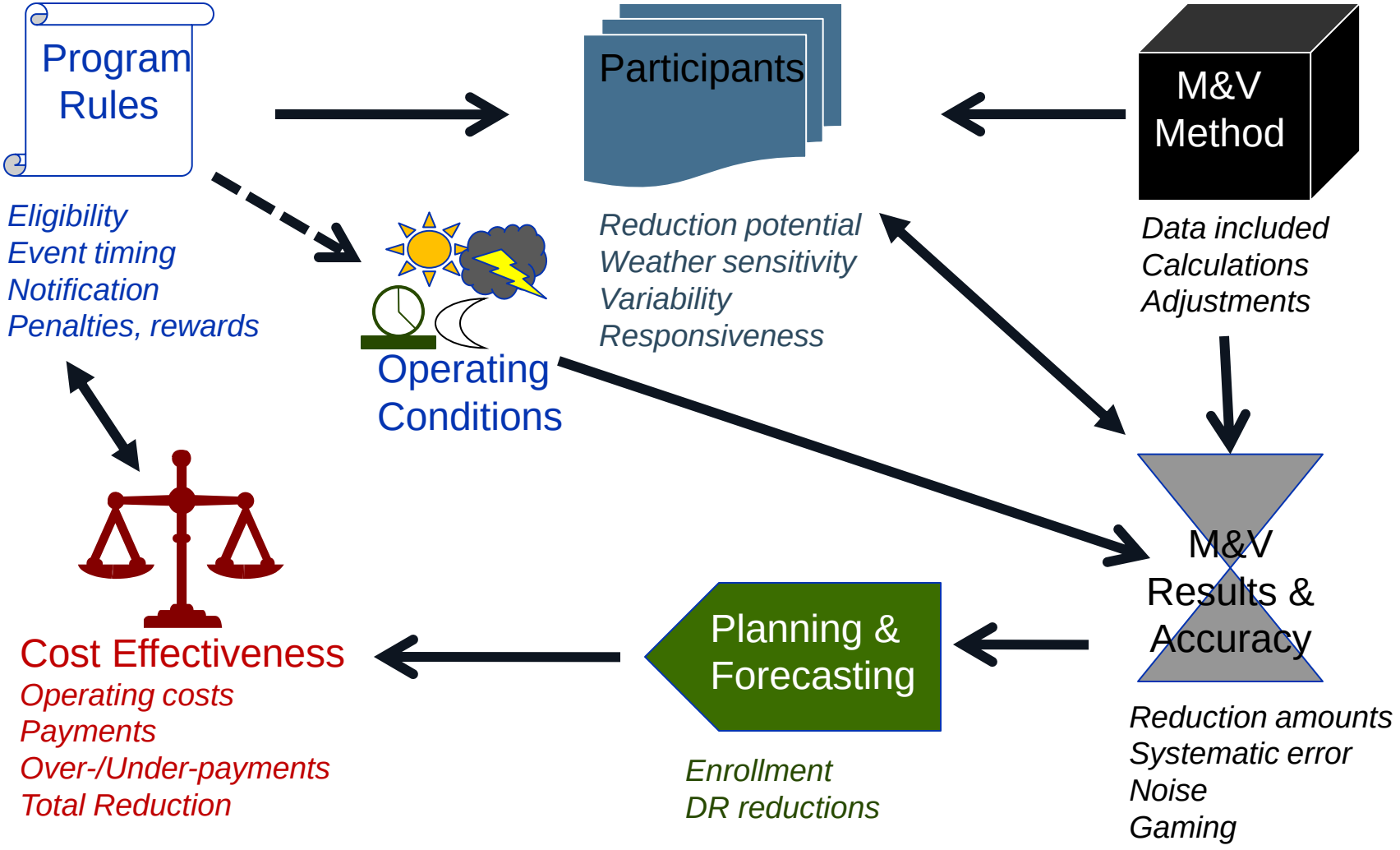
High Level Outline

- Background and goals
- Terminology and Framework:
  - NAESB Standards on DR and M&V
- Prior work
  - DR baseline method studies
  - M&V Protocols
  - Evaluation Guidelines and Protocols
- Baseline methodology issues for settlement
  - How accuracy is affected by load characteristics, program design, and M&V methods
  - Issues for particular programs
- Evaluation issues
- Recommendations and Guidance
- Gaps in our understanding
- Moving the process forward

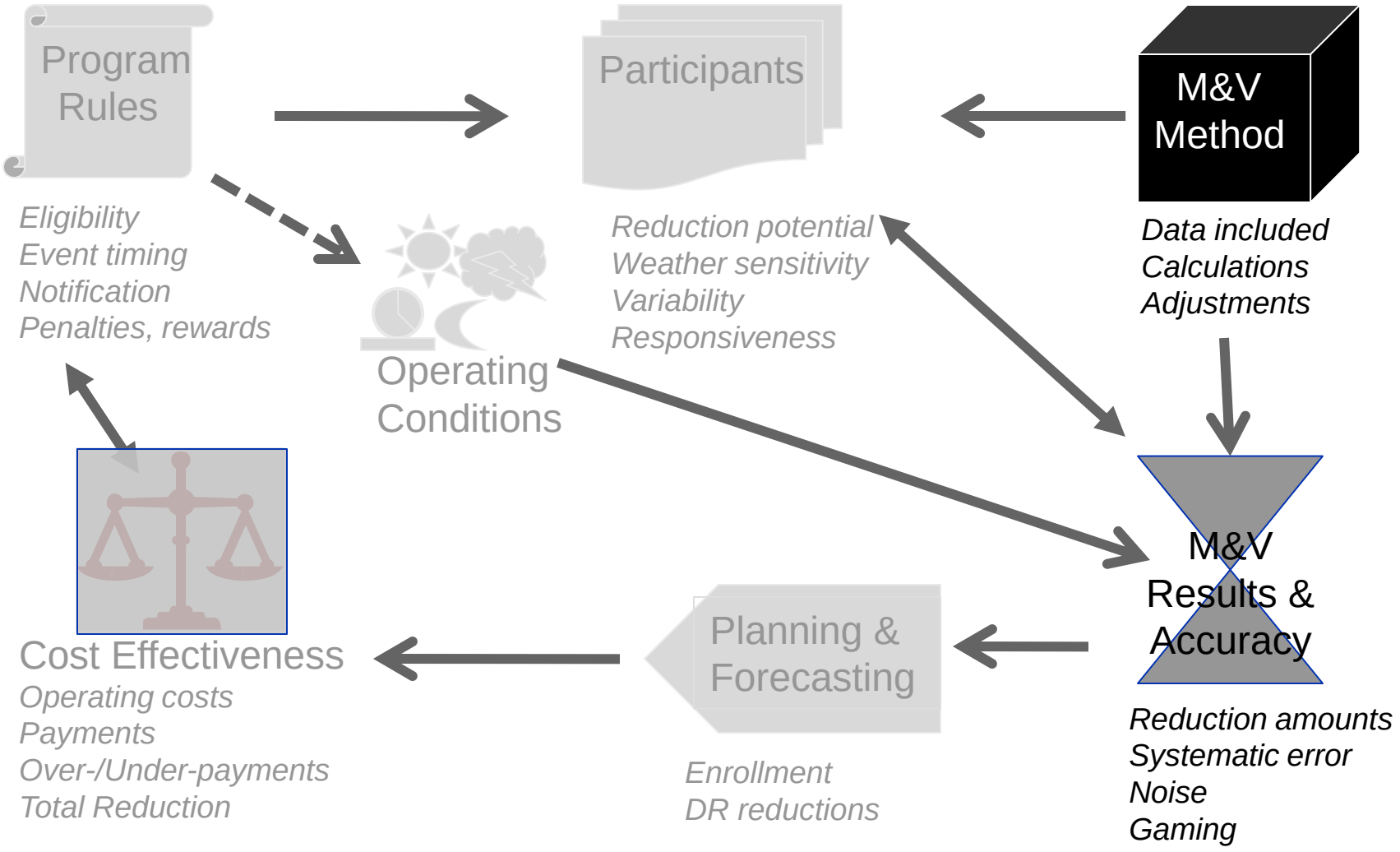
Key Guidance

- Context matters
  - Understand how the M&V will be used
  - Different methods can be useful and appropriate for wholesale settlement, retail settlement, ex post evaluation, ex ante evaluation
    - Understand how and why they're different
- All DR baseline calculation methods have errors and uncertainties
  - Understand the magnitude of bias and noise compared to true reduction quantities
    - For the loads and conditions of the given program/service/rules
- (E)M&V methods, program design, cost-effectiveness, and participating load characteristics are all inter-related.
  - Assess M&V accuracy and implications as part of program design and re-design
  - This is an ongoing process--No silver bullet and the target is always moving

DR M&V methods and results affect and are affected by many aspects of program planning, design, and operations



M&V Working Group focuses on M&V methods & accuracy, considering these contexts and relationships



Gaps Related to DR M&V (WG member suggestions, not necessarily consensus)

- Baseline-Dependent DR
  - Baseline accuracy and variability needs to be understood and accounted for as part of program design, marketing, and cost-effectiveness assessment
  - Need better understanding of baseline method performance for the residential sector
    - Most DR baseline method assessments have been for C&I sector
  - Baseline method assessments need to extend to calculating cost impacts
    - Accuracy in terms of kW is more universal
    - Impacts depend on associated prices
- General DR
  - Improve specificity of ex ante or deemed estimates used for operations
    - By time of day, weather condition
  - Increase use of experimental design for measuring program- or class-level impacts
    - Leave a randomly selected (small) subgroup uncontrolled during an event.
    - Made possible with AMI, especially for mass markets
  - Allow rules for DR that don't expect it to behave like a generator
    - Measurability, predictability

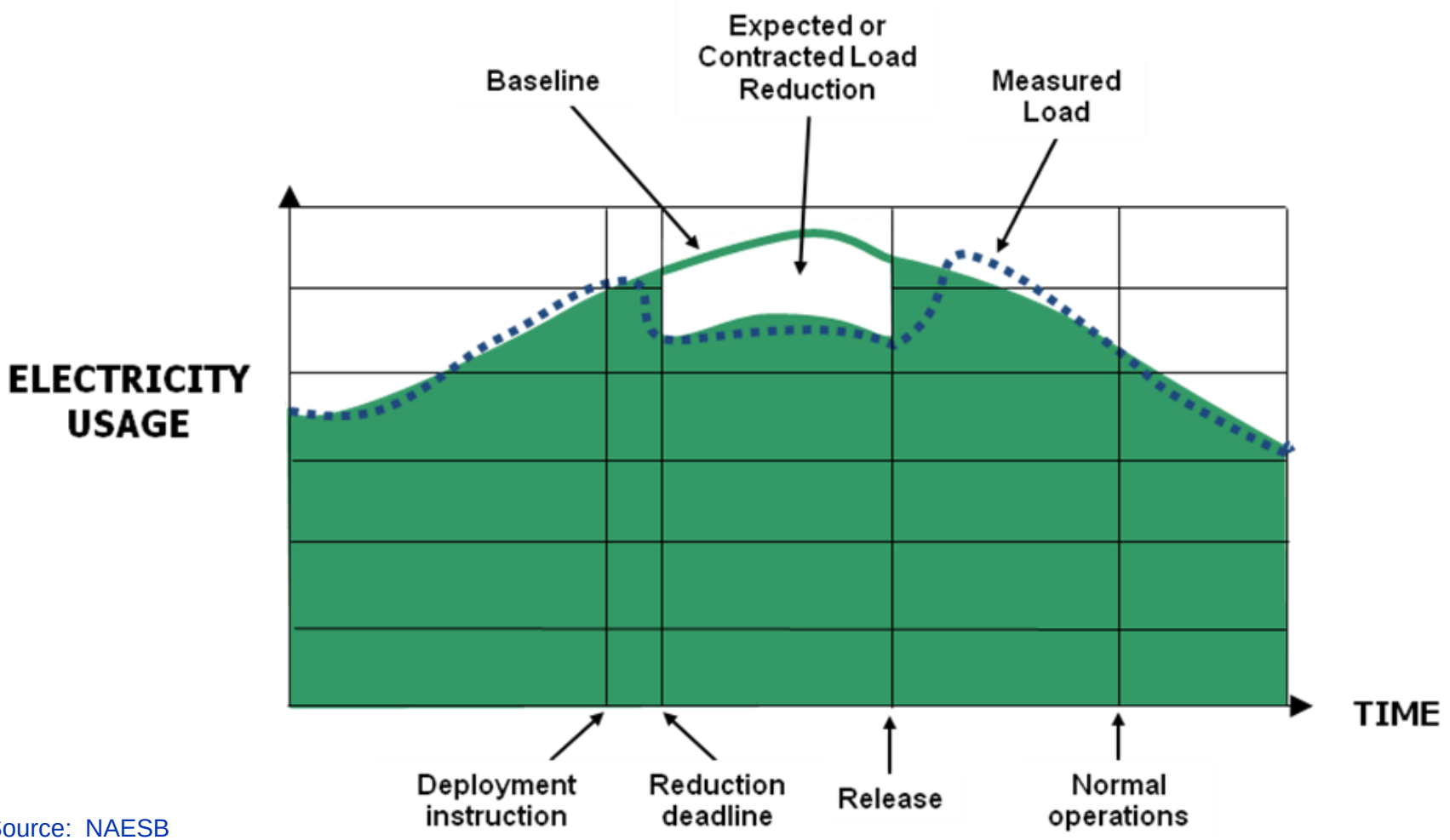
General Gaps for DR Development (WG member suggestions, not necessarily consensus)

- Demonstrate DR value and acceptance empirically
  - Ongoing process
  - Agreed cost-effectiveness framework applied in varying contexts is important
  - Meaningful best practice design criteria are important
    - Don’t decide DR doesn’t work based on programs that shouldn’t be imitated anyway
    - Learn to trust DR based on successful models

Details for Discussion

Illustration of Customer Baseline Concept

What this load would have been in these hours with no event called



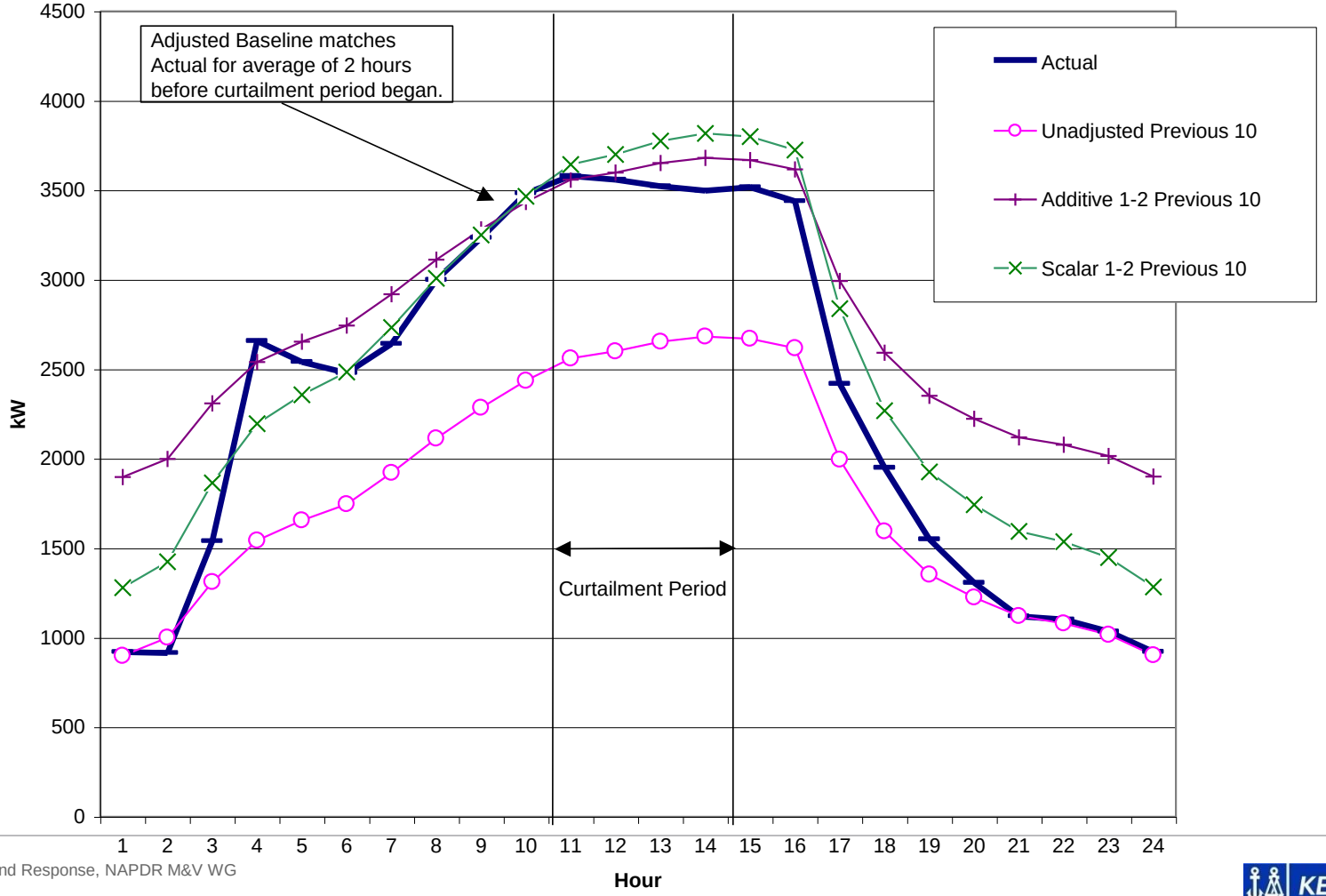
Source: NAESB
2010 WEQ AP Item 4(a) and 4(b)

NAESB DR M&V Standards

- Provide standard terminology for defining program requirements, measurement methods, and data requirements
- Identify elements that System Operators must specify for each broad type of program and performance evaluation methods
- Identify which elements and requirements are applicable to which broad types of programs (usually, unless otherwise specified by the System Operator)
- Specify firm requirements that System Operators cannot over-ride for meter accuracy and metering granularity
- Provide suggested factors that may be used for
  - Baseline calculation types
  - Data exclusion rules
  - Baseline Adjustments
- Retail standards include high level requirements for statistical sampling methods
- Not included
  - Guidance on specifications for particular program/market rules and resource characteristics

Illustration of Adjustments

Unadjusted baseline: Average or regression from recent history per data selection rules



What we understand about customer baselines for DR settlement

- For large C&I, many simple baseline methods for individual loads can have systematic errors under 2% and random noise under 15% for a wide variety of loads and conditions
  - These errors can still be large as percent of *reduction*
- Weather-sensitive loads
  - It's hard to construct accurate baseline if events are called/bids clear on all hot days
  - Without day-of-event adjustment, reductions on hot days can be substantially understated
- Day-of-event adjustments
  - Are typically necessary to ensure reward/incentive for weather-sensitive loads on hot days
  - Tend to over-state reduction for customers who pre-cool in response to notification or in anticipation of likely event.
  - Tend to understate reduction for customers who cancel shifts before event in response to notification or in anticipation of likely event
  - Should be based on customer's own load pre-notification, or on system/weather factors
- Highly variable loads (apart from weather)
  - For these loads, general customer baseline methods tend to produce DR calculations almost unrelated to actual DR actions
    - Potentially discourages their participation
  - Consider tailored baseline calculations for large high-variability loads
    - Shift prediction burden to the customer

Potential Adjusted Baselines for Settlement

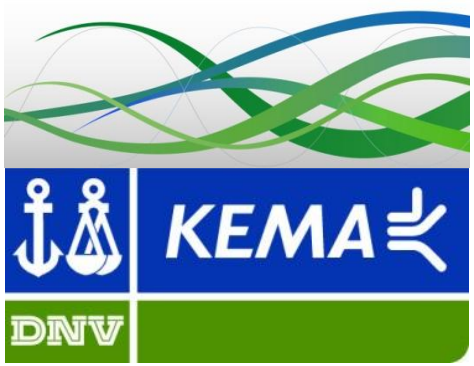
If Notification Is	For Load Characteristics: Variability (apart from weather)	For Load Characteristics: Weather-Sensitivity	A Useful Adjustment Basis is	Likely Accuracy Problems after Adjustment are
Same day	Low	Low	None or Own load, 1-2 hrs pre-notif'n	Minimal
	Low	High	Own load, 1-2 hrs pre-notif'n	Precautionary pre-cooling can inflate baseline
	High	Low	Own load, 1-2 hrs pre-notif'n	Underlying unpredictability
	High	High	Own load, 1-2 hrs pre-notif'n	Precautionary load shifts can inflate baseline Underlying unpredictability
Day ahead	Low	Low	None	Minimal
	Low	High	System or weather, 1-2 hrs pre-notif'n	Pre-cooling in response to notification inflates baseline
	High	Low	System or weather, 1-2 hrs pre-notif'n	Underlying unpredictability
	High	High	System or weather, 1-2 hrs pre-notif'n	Preparatory load shifts inflate baseline Underlying unpredictability

Concerns about participants manipulating their baselines (Behavior designed to inflate customer baselines)

- No baseline calculation method can eliminate possibility of manipulation
- Manipulation or “gaming” doesn’t happen unless it’s worth the trouble
  - Energy and operational/inconvenience costs have to be less than expected excess payment
  - Likely to be an issue mainly for some C&I loads
  - Most participants have better things to do
- Reducing opportunities for baseline manipulation by participants
  - Start with a method that’s fair on average on likely event days, absent any gaming
  - Ensure baseline calculation data include recent “similar” days
  - Limit participants’ ability to control or predict what days they will be called on to reduce
  - Investigate load and bidding patterns that seem perverse based on customer type
  - Require advance notice of scheduled shut-downs
- Preparatory load shifting isn’t necessarily gaming or manipulation
  - Pre-cooling after event notification, before event start
  - Pre-cooling on a forecasted hot day before event notification
  - Shifting manufacturing loads to overnight in anticipation of next day event

Impact Evaluation

- Impact evaluation measures overall program demand reductions
  - In total and by customer segment
  - For individual events *ex post* and for particular conditions *ex ante*
  - Typically not concerned with accuracy for individual customers, but with aggregate accuracy
- Uses include
  - Program planning and re-design
  - Short and long-term operational forecasts
  - Settlement with ISO for retail program
- Methods
  - Not limited to totaling settlement amounts over individual customers
  - Can use regression models with many parameters
  - Can use data from full season or multiple seasons
  - Can fit pooled models across customers
  - Should include statistical accuracy measures where applicable
- Evaluation results can demonstrate DR as a reliable, measurable resource
  - Even if individual customer baselines have noise and known biases



Contact Information:

Miriam Goldberg
Senior Vice President
DNV KEMA Energy and Sustainability
Miriam.Goldberg@dnvkema.com
608-259-9152