MAXIMIZATION OF THE VALUE OF INTRA-DAY WIND AND SOLAR FORECASTS FOR AN ISLAND GRID SYSTEM VIA CUSTOMIZED FORECASTS AND EVALUATION METRICS

John W Zack, Ph.D.
UL Renewables
463 New Karner Rd
Albany, NY
jzack@awstruepower.com

Robert Kaneshiro and Lisa Dangelmaier
Hawaii Electric Light Company (HELCO)
Hilo, Hawaii
robert.kaneshiro@hawaiielectriclight.com

OUTLINE

1. Objective and Approach
   - Key Issue: forecast value is NOT equivalent to forecast accuracy
2. Background on the Target System - Big Island, Hawaii
3. Conceptual Forecast Evaluation Approach and Metrics
4. Three Key Types of Operations Decisions
   - Decision factors and implications
   - Key forecast attributes for decisions
   - Application-relevant metrics
5. Summary and Next Steps

OBJECTIVE AND APPROACH

- **Objective**: Maximize the value of wind and solar generation forecast information for operational decision-making
- Why is this an issue?
  - o Most users employ forecasts that are not optimized for their applications
  - o Forecast performance is typically measured with generic metrics (e.g. MAE, RMSE) that are often not strongly related to the way in which the user's applications are sensitive to forecast error
  - o Most users have not determined which attributes of forecast performance are most important for their applications
- **Result:** A considerable amount of forecast value is unrealized (or from another perspective: thrown in the trash)

- 4-part project to maximize forecast value for HELCO has been initiated

1. Define key decision-making time frames and issues (scenarios)
2. Formulate parameters that quantify the key forecast issues in each scenario
3. Evaluate forecasts of these parameters with application-customized metrics
4. Optimize forecasts to optimize performance on application-customized metrics

Addressed in this presentation

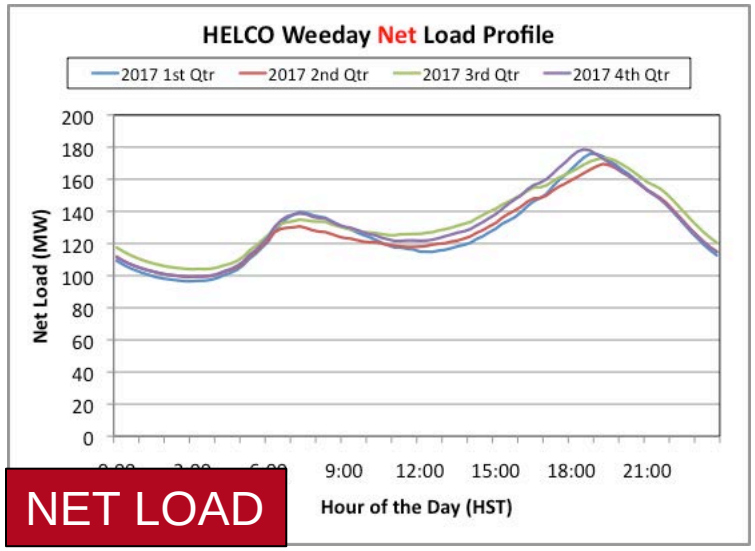
BACKGROUND INFORMATION

HELCO Load and PV Profiles

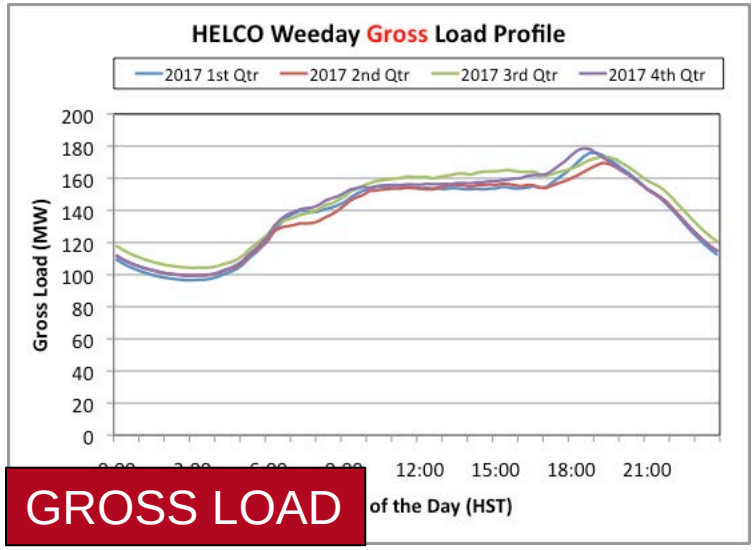
HELCO Generation Resources

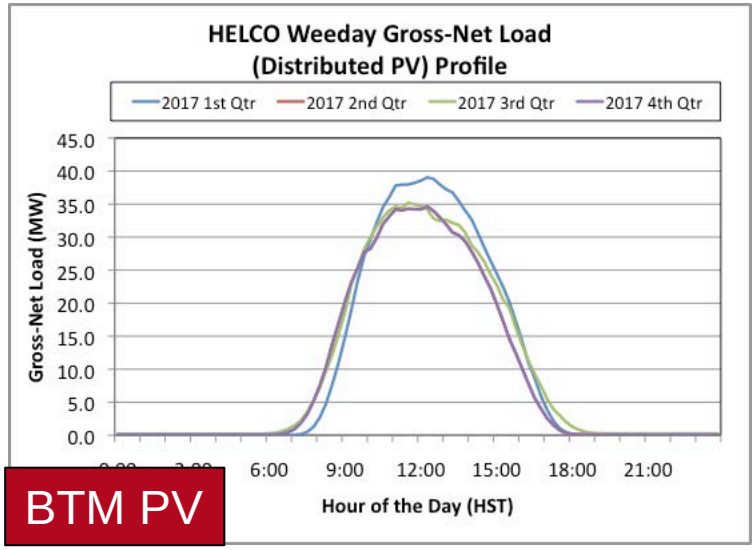
SWIFT Forecast Specifications

HELCO: 2017 WEEKDAY LOAD PROFILES

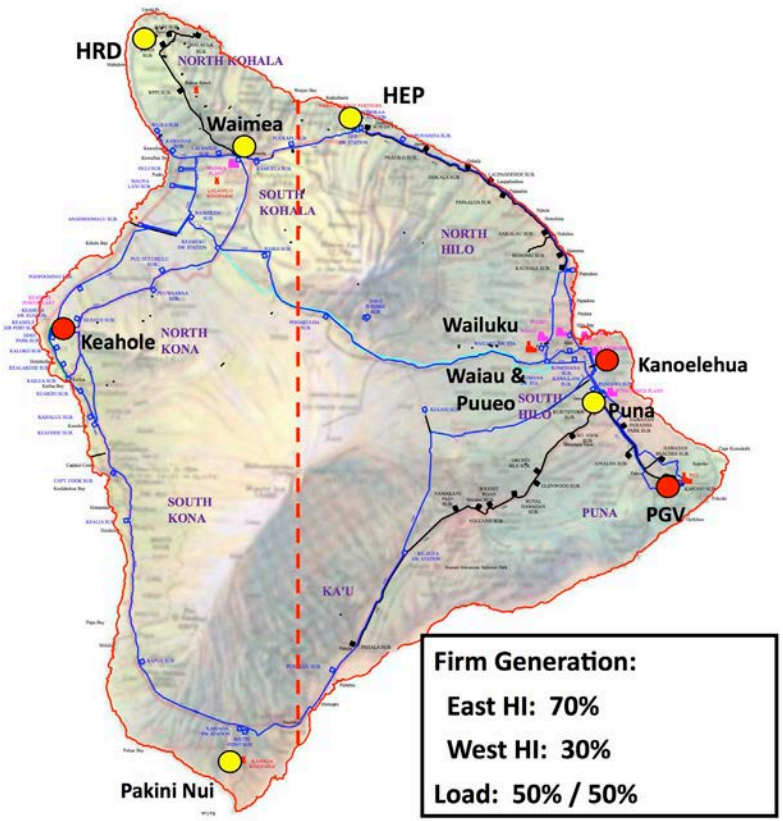


- Weekday Net load: 2 daily peaks
  - Morning (~0800): 130-140 MW
    - Morning rise in gross load followed by morning rise in PV production
  - Evening (~1800): 170-180 MW
- Weekday Net load: 2 daily minima
  - Nighttime (~0300): 95-105 MW
  - Daytime (~1200): 115-125 MW
    - Associated with peak of 35-40 MW of distributed PV production





HELCO: GENERATION RESOURCES - 2017

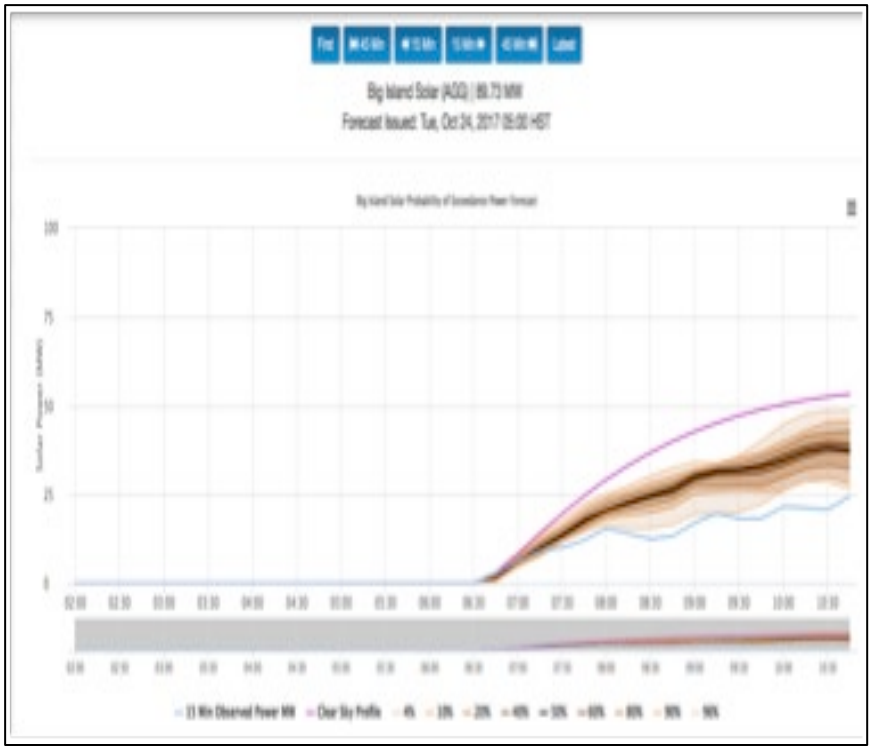


Base (24-hr) Units
Hill 5 & 6 Steam Units
Keahole 1CT in combine cycle (CC)
PGV (Geothermal)
Intermediate Units
Keahole 2nd unit in CC
HEP 1st and 2nd in CC
Peaking/Emergency Units
Kanoelehua CT-1
Keahole CT-2
Puna CT-3
Puna Steam Unit
12-Small Diesel Generators
As-Available Must-Take
HRD Wind farm (10.5 MW)
Pakini Nui Windfarm (20.5 MW)
Wailuku River Hydro (12 MW)
Puueo Hydro (3.1 MW)
Waiau Hydro (1.1 MW)

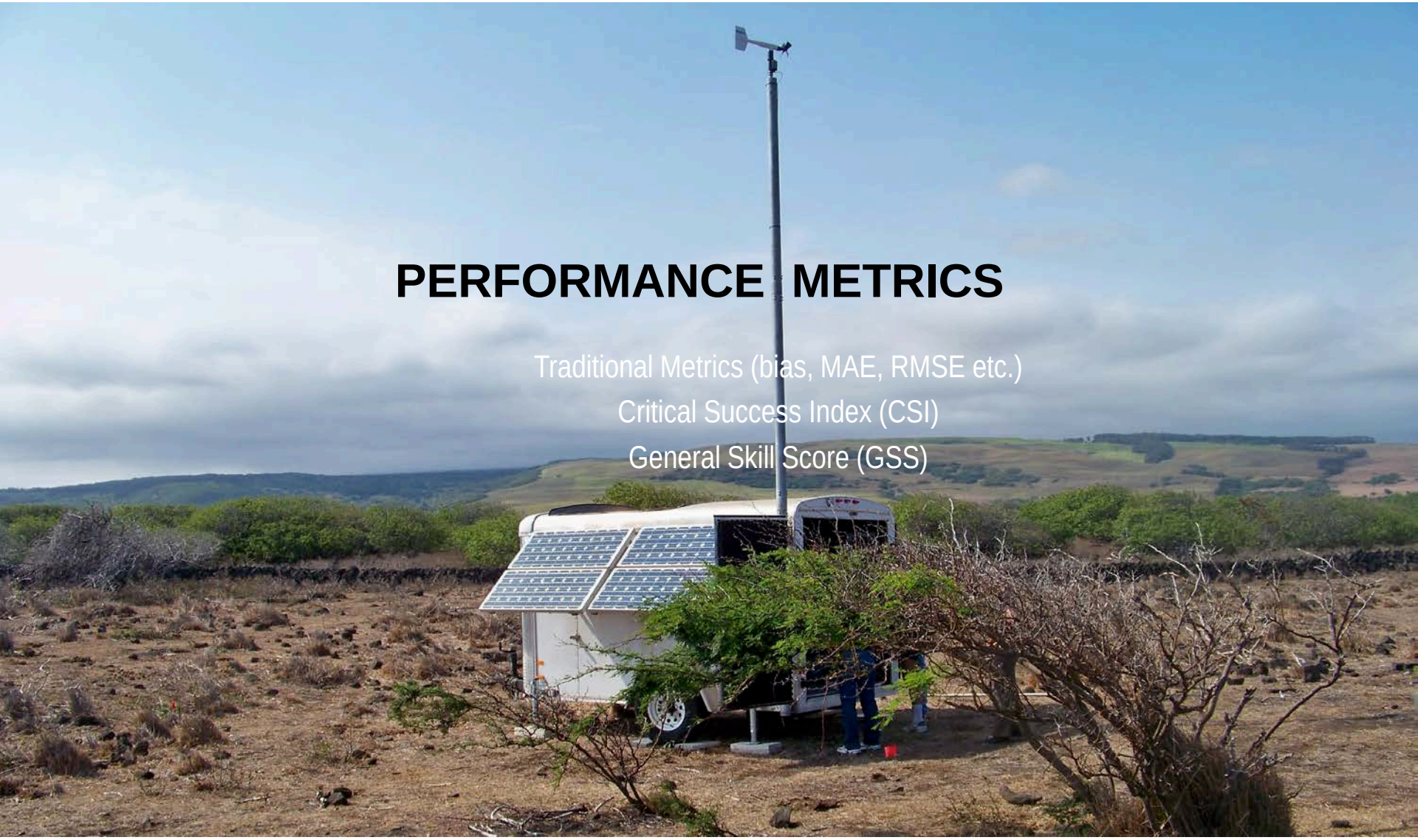
Type of Resource	Capacity
Geothermal	38 MW
Hydro (3 facilities)	16.2 MW
Wind (2 facilities)	31 MW
Solar (distributed behind the meter)	90 MW

WIND AND SOLAR FORECAST SYSTEM: SWIFT

Example of a SWIFT Forecast Display



- Forecasts are produced from an ensemble of prediction-methods (physics-based and statistical)
- Two Forecast Time Frames
  - Intra-day
    - 0-6 hrs ahead in 15-min time steps
    - 15-min updates
  - Multiple Day
    - 0-7 days ahead in 1-hr time steps
    - 1-hr updates
- Probabilistic Format
  - 10 Probability of Exceedance (POE) values
  - 50% POE used as deterministic forecast
- Target Entities
  - Utility-scale PV & wind generation facilities
  - Substation aggregates of distributed PV
  - Regional and system (island) aggregates



PERFORMANCE METRICS

Traditional Metrics (bias, MAE, RMSE etc.)

Critical Success Index (CSI)

General Skill Score (GSS)

TRADITIONAL FORECAST PERFORMANCE METRICS

- Mean Error (bias)
  - Average of the errors over all forecast intervals in a sample
  - Provides an indication of the systematic error (e.g. too low or too high)
- Mean Absolute Error (MAE)
  - Average of the absolute value of the errors over all forecast intervals in a sample
  - Provides an indication of the typical magnitude of the error
- Root Mean Square Error (RMSE)
  - Square root of the average of the squared errors
  - Provide an indication of typical error magnitude but heavier weighting of larger errors

- Issues
  - Heavily weighted towards performance under typical conditions
  - Not sensitive to performance under atypical (often the most critical) conditions
  - Not focused on events that are most important to decision-making

CONCEPTUAL APPROACH

- HYPOTHESIS: A customized event-oriented category-based evaluation scheme would provide a more application-relevant assessment of forecast performance than a traditional forecast evaluation approach
  - Events defined by critical operational decision-making scenarios
  - Categories defined by operationally significant thresholds associated with the events
  - Events, categories and time frames customized for each decision-making scenario
  - Category-based performance metrics used to assess forecast performance

CATEGORY-BASED EVALUATION: CSI

Category #	Forecasted			
Onserved	Category	Below	Typical	Above
	Below	1	2	3
	Typical	4	5	6
	Above	7	8	9

- Ratios of Correct and Incorrect Outcomes
  - Hits (H) = Cat #1 + Cat #9
  - Misses (M) = Cat #2 + Cat #3 + Cat #7 + Cat #8
  - False Alarms (FA) = Cat #4 + Cat #7 + Cat #3 + Cat #6
  - Critical Success Index (CSI) = H / (H + M + FA)
- Issues
  - Does not account for multiple category errors
  - Does not consider relative frequency of outcomes: could provide hedging incentive
  - Does not weight relative cost of errors
    - Does a miss of a “below” event cost the same as a miss of an “above” event?
    - Is the cost of a “miss” the same as a “false alarm”?

CATEGORY-BASED EVALUATION: GSS

- General Skill Score (GSS)
  - Measures skill relative to a random forecast of categories considering the relative frequencies of outcomes (0= same as a random forecast, 1= perfect)
  - Can be formulated to have relative weighting for errors
  - All of this accomplished through a scoring matrix: s_{ij}

$$GS = \frac{1}{N}\sum_{i=1}^{K}\sum_{j=1}^{K} n(F_i, O_j)\, s_{ij}$$

N = Total # of fcst-outcome pairs
n(F,O) = # of pairs in each fcst-outcome bin
S = Scoring matrix (score for each bin)
K = # of forecast categories

- Example of a Scoring Matrix (s_{ij})
  - Based on a 10%, 80%, 10% (below, typical, above) frequency of outcomes
  - 2-category errors are penalized twice as much as a 1-category errors
  - All other errors have the same weighting (misses, false alarms etc.)

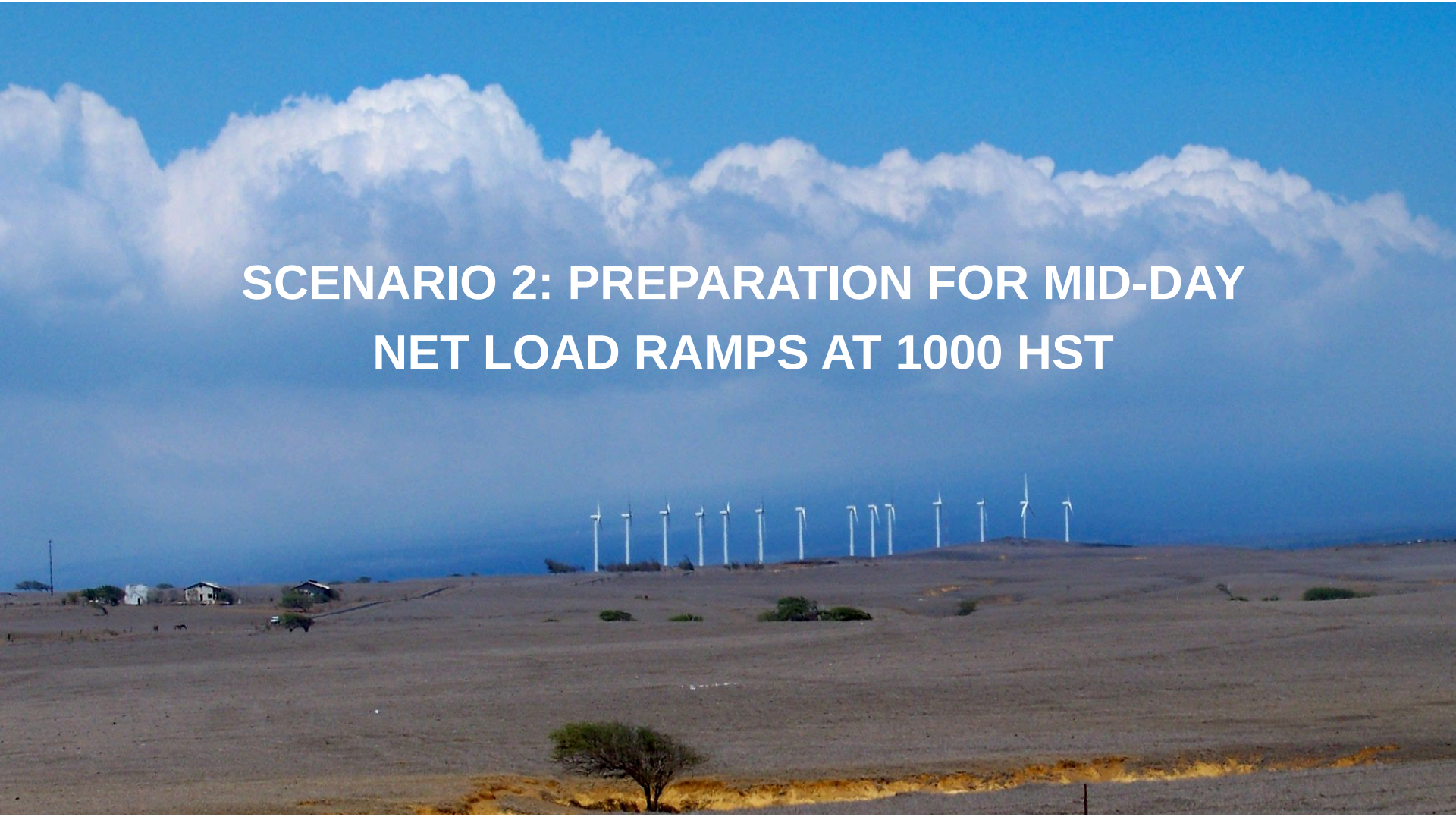
SCORING	Forecasted			
Observed	Category	Below	Typical	Above
	Below	1	-4	-1
	Typical	0	1	0
	Above	-1	-4	1

KEY DECISION MAKING TIMES AND ISSUES

KEY DECISION-MAKING TIME FRAMES AND ISSUES

- Three key daily time frames and issues were identified
  - Scenario 1: 0500 HST Preparation for morning peak and mid-day minimum
  - Scenario 2: 1000 HST Midday net load ramps
  - Scenario 3: 1300 HST Preparation for evening peak
- Each are sensitive to different aspects of forecast performance

Focus of this presentation

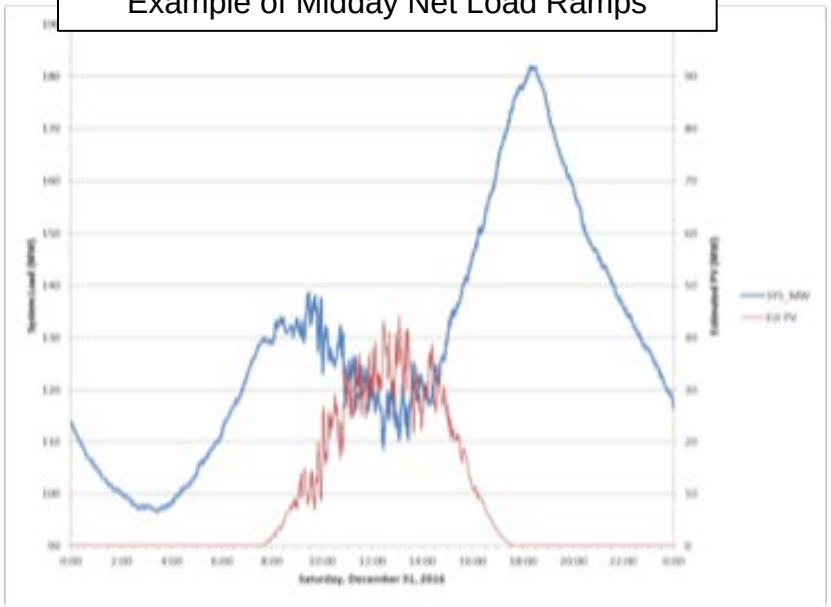


SCENARIO 2: PREPARATION FOR MID-DAY NET LOAD RAMPS AT 1000 HST

LATE MORNING (1000 HST) DECISION-MAKING PERIOD

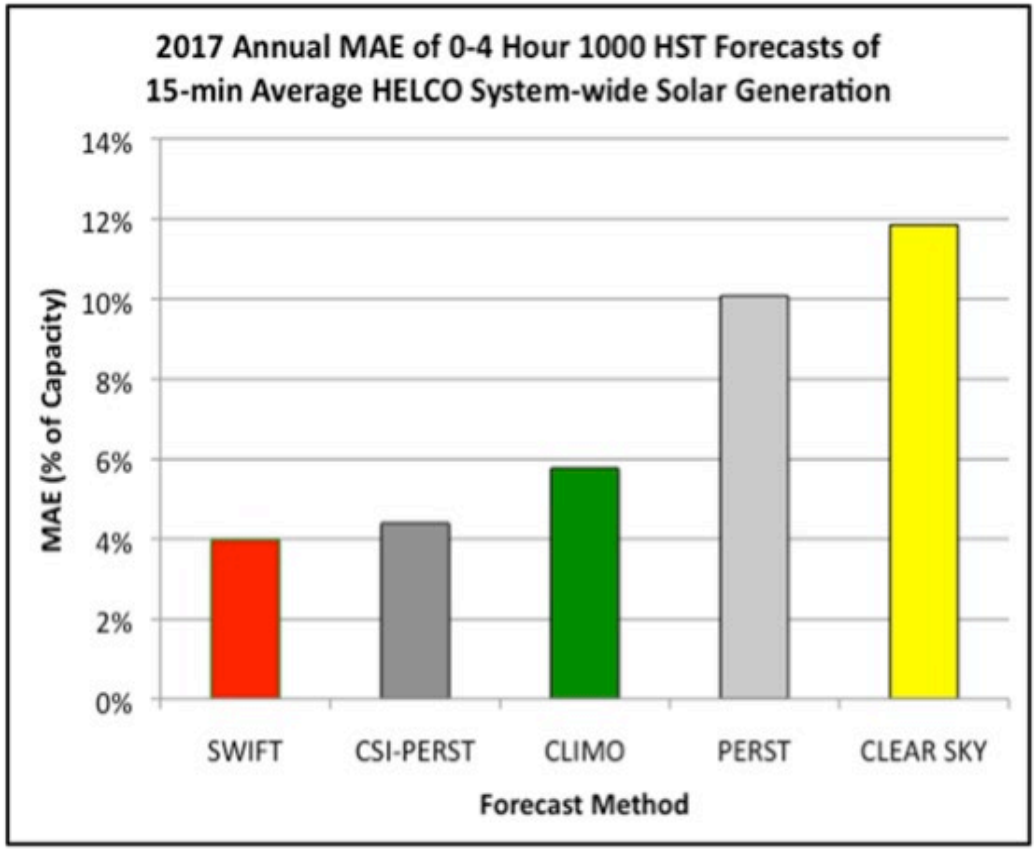
- System Management Issues
  - Will significant mid-day net load ramps occur due to large variability in BTM PV production?
- Critical Forecast Questions
  - Will short-time-scale (30 minutes or less) variability be above average?
  - What will be the amplitude of the short-time scale variability?

Example of Midday Net Load Ramps

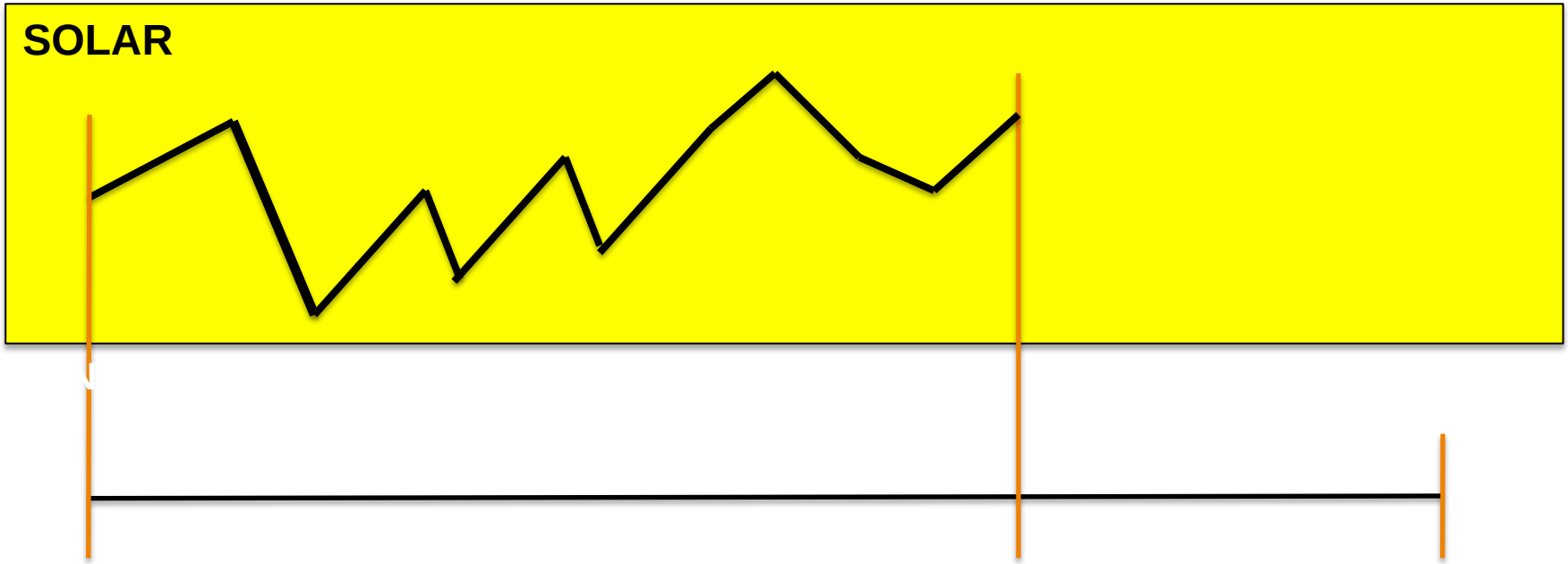


MID-DAY SOLAR VARIABILITY:
1000 HST FORECAST: TRADITIONAL EVALUATION

- Metric: Mean Absolute Error (MAE)
- SWIFT and 4 Reference Forecasts
  - CSI-PERST: no change in clear sky generation index from forecast issue time
  - CLIMO: long-term average by time of day and month
  - PERST: no change in solar gen from forecast issue time
  - CLEAR SKY: solar gen if no clouds
- Conclusion: SWIFT midday solar forecasts perform well in a MAE-based evaluation



PARAMETERS THAT MEASURE MID-DAY VARIABILITY



1000 HST: Forecast Issue Time

1400 HST

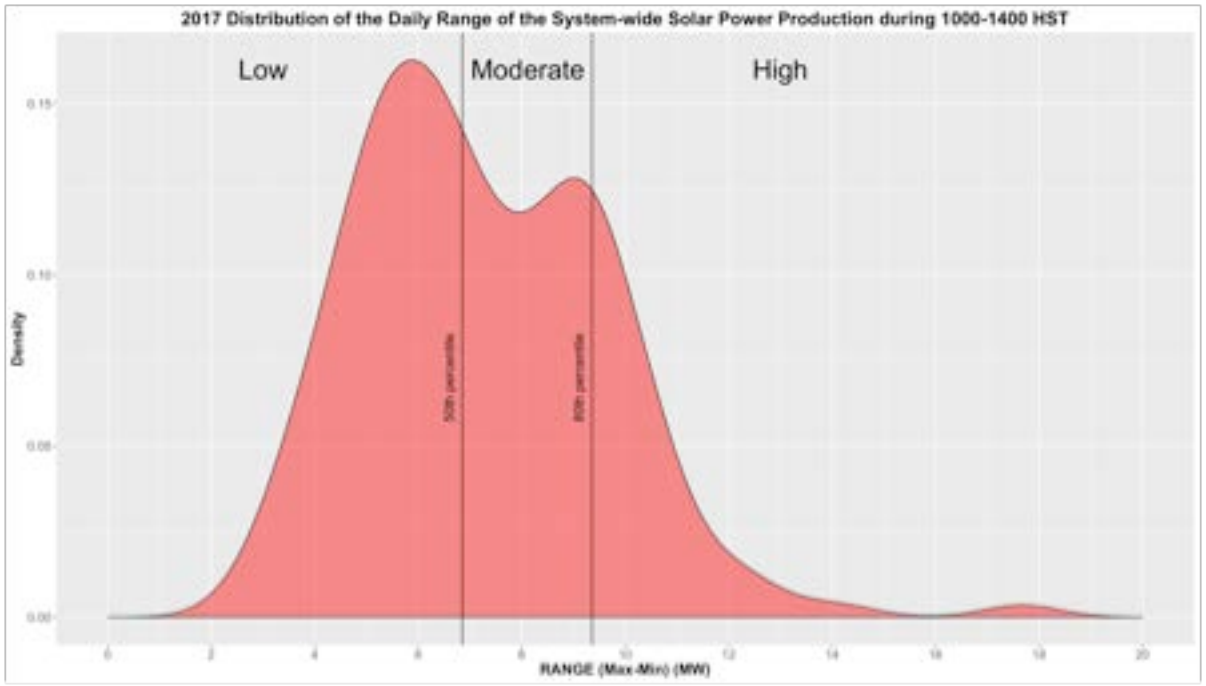
1600 HST

- Four possibilities considered:

1. Max 15-minute ramp rate
2. Standard deviation of 15-minute averages
3. Total 4-hr path length (sum of abs value of 15-min changes)
4. Range (maximum – minimum of 15-min averages)

MID-DAY SOLAR VARIABILITY: CATEGORY-BASED EVALUATION

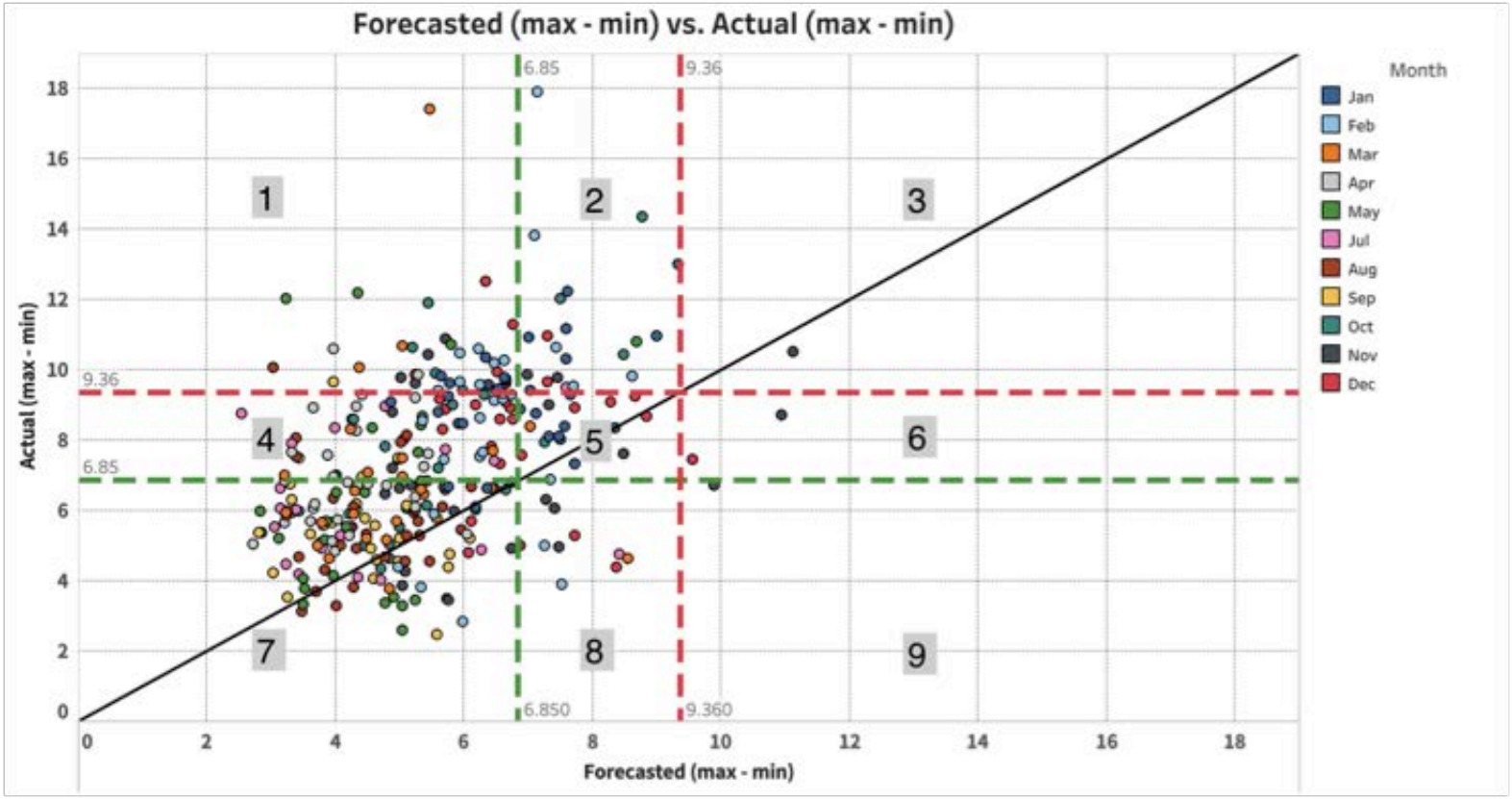
- Variability measured by the RANGE parameter: MAX-MIN of 15-min average values
- Define 3 categories of variability for the 1000 – 1400 HST period based on 2017 data
  - o Low: RANGE below 50th percentile
  - o Moderate: RANGE between 50th and 80th percentile
  - o High: RANGE above 80th percentile



MID-DAY SOLAR VARIABILITY:
1000 HST FORECAST VS. ACTUAL RANGE

2017 Daily Forecast vs. Actual 10-14 HST Range (Max-Min) in System-wide 15-min Average PV Generation

CONTINGENCY		Forecasted			
	Category	F_1:Low	F_2:Moderate	F_3:High	Obs %
Observed	O_3:High	40	21	1	20.0%
	O_2:Moderate	72	20	2	30.3%
	O_1:Low	143	10	1	49.7%
	Forecast %	82.3%	16.5%	1.3%	100.0%



MID-DAY SOLAR VARIABILITY:

RANGE FORECAST PERFORMANCE METRICS

CSI and Its Components for 2017: 2 Event Definitions

EVENTS	H	M	FA	Total	CSI
Moderate or High (>50th Percentile)	44	112	11	167	26.3%
High (>80th Percentile)	1	61	3	65	1.5%
Composite (All Events)	45	173	14	232	19.4%

General Skill Score (GSS) for 2017

GSS		Forecasted		
Observed	Category	F_1:Low	F_2:Moderate	F_3:High
	O_3:High	-74.3	-9.0	1.0
	O_2:Moderate	-30.9	20.0	0.5
	O_1:Low	143.0	-4.3	-0.5
Category Scores		37.9	6.7	0.9
Total Score				45.5
Number of Forecasts				310
GSS (Total Score/# Forecasts)				14.7%

SUMMARY AND NEXT STEPS

SUMMARY AND NEXT STEPS

- Key operational decisions and time frames that are impacted by forecast information and associated critical forecast performance attributes were identified
- A flexible category-based forecast evaluation scheme focused on events that have a significant impact on operational decision-making was developed
- A comparison of 2017 forecast performance from the traditional and application-customized evaluation schemes was compiled

Forecast Issue Time	Generation Type	MAE	CSI	GSS
0500 HST	Solar	4.22%	37.3%	39.1%
0500 HST	Wind	13.48%	5.0%	6.5%
1000 HST	Solar	3.98%	13.2%	14.7%
1300 HST	Solar	3.69%	29.5%	31.7%
1300 HST	Wind	10.67%	4.8%	-1.2%

- Next Steps
  - Customize forecast systems to optimize performance for critical forecast attributes
  - Assess performance of customized forecasts with application-relevant metrics