SSL Statistical Analysis of Models Related to TM-21

May 11th, 2010 -- CORM 2010 Technical Conference

Trenton Pulsipher, John Hathaway, Amy Qiao, Eric Richman

Pacific Northwest
NATIONAL LABORATORY

TM-21 and the Analysis Process

- What is TM-21?
  - An IES Technical Memorandum - Intended to be a companion to LM-80 (LED lumen degradation measurement method)
  - To provide a method of estimating or predicting lumen maintenance/degradation as a part of LED product life.
  - Applies an estimation method to the 1,000 hour lumen degradation measurements over 6,000 hours (specified by LM-80)
- Analysis process
  - Explore model characteristics of data fit and true representation of possible decay trends.
  - Evaluate reasonable models with sample and real data to explore model choosing methods.
  - Use results to determine a reasonable degradation estimate method.

Data and Modeling

- LED Lumen Degradation Model Development
  - **Engineering Derived Models:** LED experts must identify the decay effects which model the lumen degradation of an LED.
  - **Statistically Derived Models:** Requires complete data over the life or representative LEDs.
    - 6,000 hours of data is not enough data to build a reasonable statistics based model.

Proposed LED Lumen Degradation Models (Engineering)

- Originally proposed models for LED lumen degradation
  - Emil Radkov (Illumitex) developed the following models.
  - It may be advantageous to create other models, but this work was restricted to statistical estimation using these models.

Model	Decay Rate	Comments
0	$dI_v/dt = 0$	Data Mean
1	$dI_v/dt = k_1$	Linear
2	$dI_v/dt = k_2 I_v$	Simple Exponential
3	$dI_v/dt = k_1 + k_2 I_v$	Linear + Exponential
4	$dI_v/dt = k_3/t$	Log
5	$dI_v/dt = k_1 + k_3/t$	Linear + Log
6	$dI_v/dt = k_2 I_v + k_3/t$	Exponential + Log
7	$dI_v/dt = k_1 + k_2 I_v + k_3/t$	Linear + Exponential + Log

Data and Modeling Characteristics

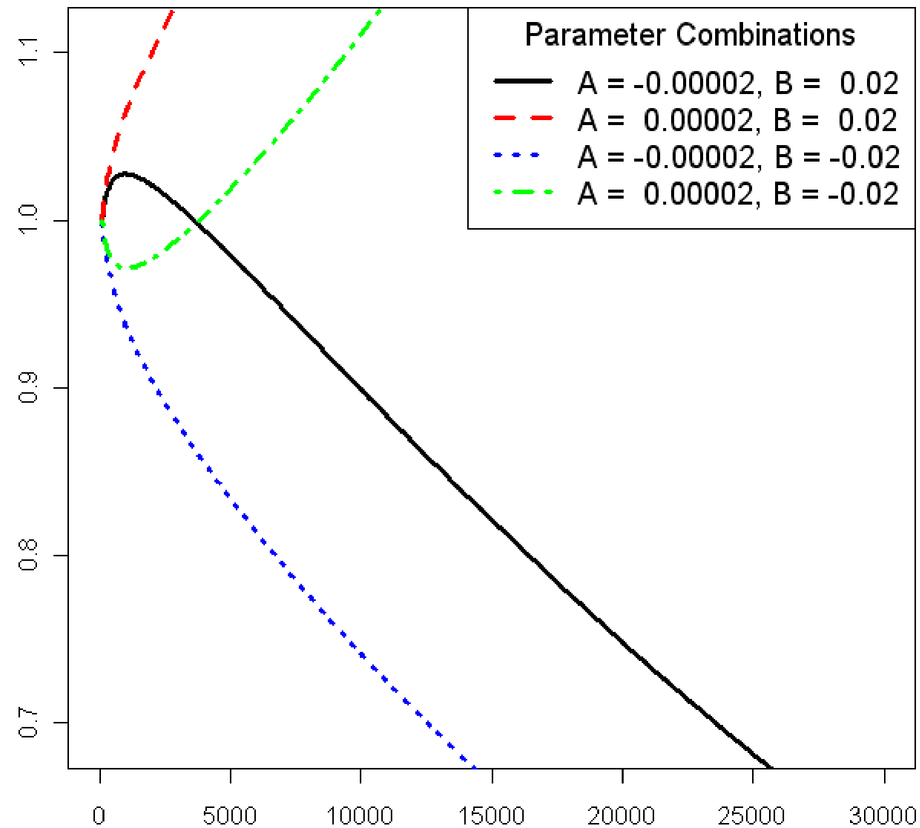
- **Model selection:** The Sum of Squares Error (SSE) values of multiple models are often too close to make reliable selection of a single model.
- **Parameter estimation (slope, intercept, etc.):** Still need estimates of the model parameters even if the correct model is known.
  - With limited data (first 6,000 hours): parameter estimation is difficult.
  - Models with more parameters require more data for accurate estimation.
- **Multiple parameters vs. less parameters:** Models with multiple parameters have capability to describe complex decay structures **but are more sensitive to data noises, therefore, tend to over fit simple decay structures.**

Models 6 (Exponential + Log) & Model 7 (Linear + Exponential + Log) Characteristics

Model Fitting Issues

- **Excessively Complex:** These models fit in the derivative space (not the data space) so have potentially large fit errors.
- **Model Structure vs. Data Structure:** Compared to the limited data (first 6,000 hours) the model structure is very complicated – built to model a hump or initial data structure near time 0.
- **Model 6** – The first point is highly influential due to derivative structure
- **Model 7** – Has the most parameters to estimate causing potential fit errors

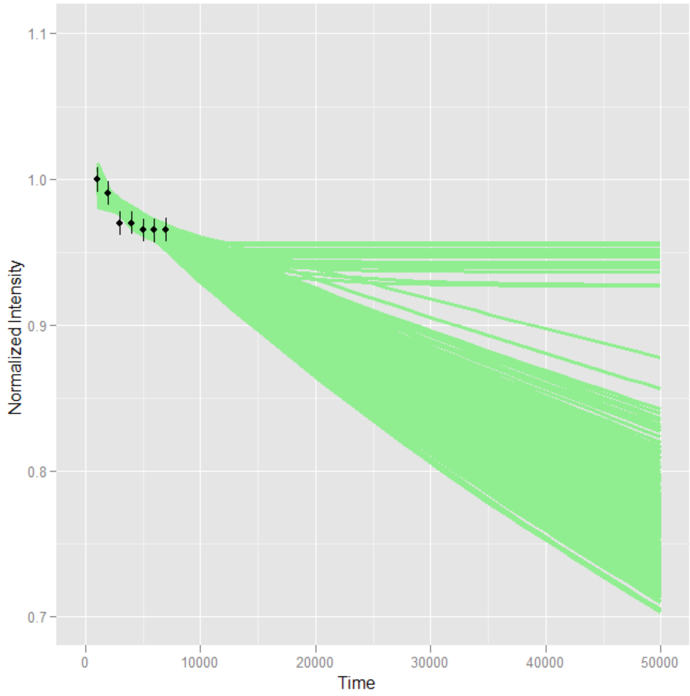
Examples of Possible Model 6 Curves

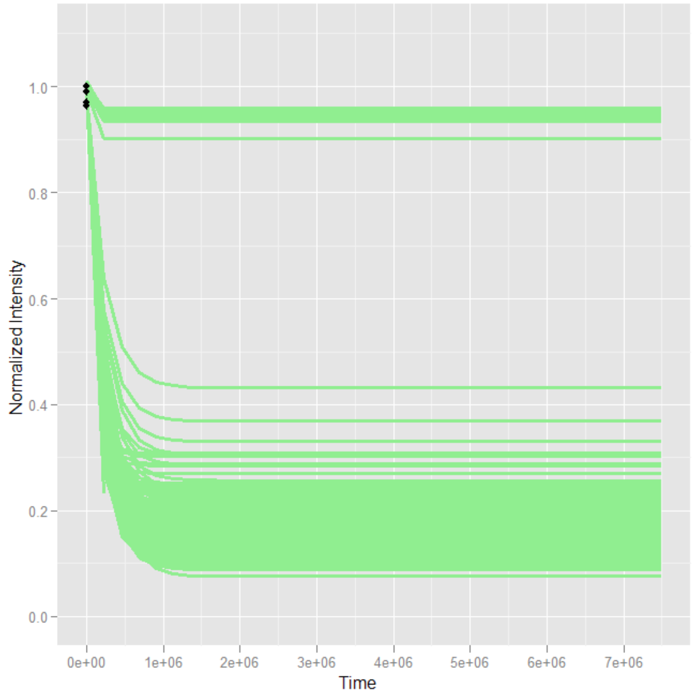


Pacific Northwest NATIONAL LABORATORY

Model 3 (Linear + Exponential) Characteristics

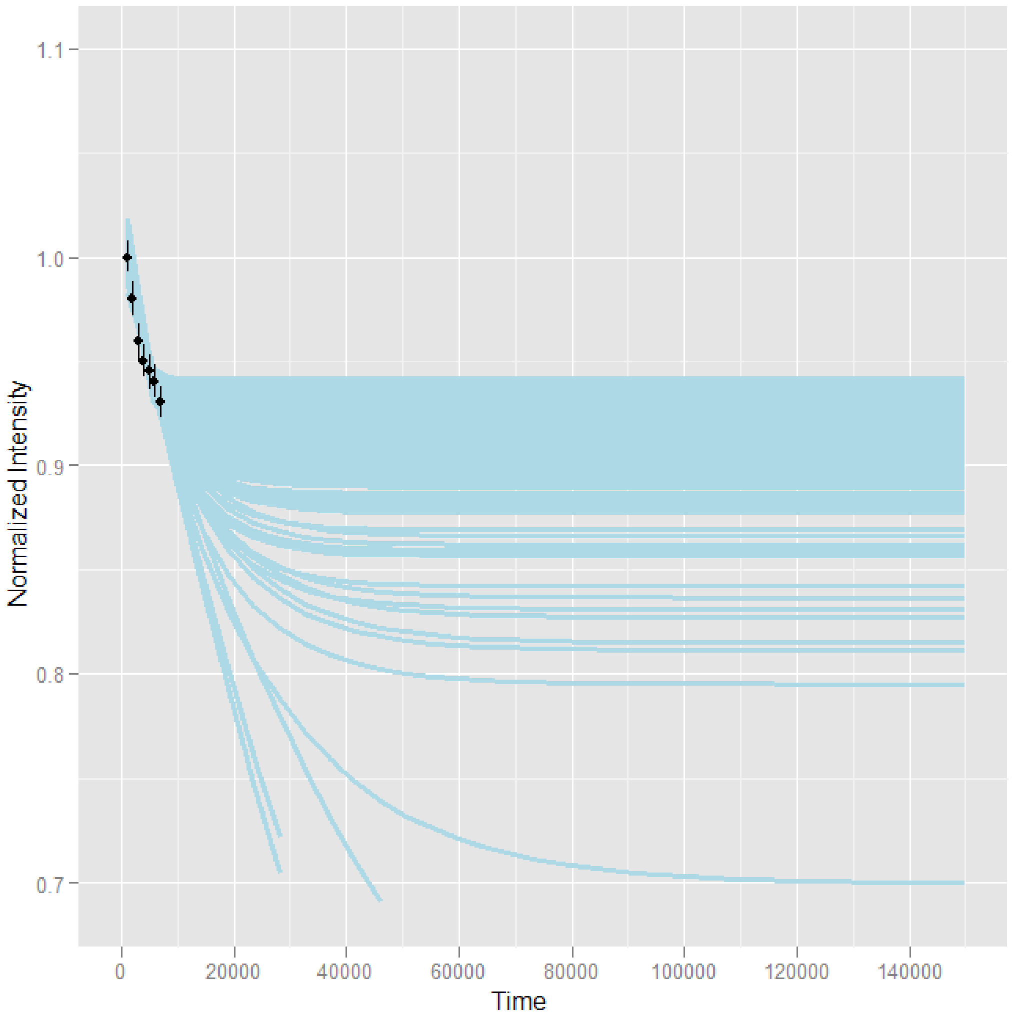
- Model 3 (Linear + Exponential) asymptotes and results in infinite L_{70} predictions
  - **Model Structure vs. Data Structure:** The data structure provides little evidence for the predicted model structure.
  - **Noisy Data:** Overly sensitive to noise or slight variations in the final measurements (e.g. 5,000 and 6,000 hours).





Model 4 (Log) Characteristics

- Model 4 (Log) often suggests L_{70} predictions of several hundred years
  - **Model Structure vs. Data Structure:** The model structure is too rigid /set for the limited data structure.
  - These “unbelievable” L_{70} predictions are inappropriate considering the number of observations.



Pacific Northwest NATIONAL LABORATORY

Summary of Model Characteristics Exploration

Evaluated Models (blue) & Not Evaluated Models (black)

- Model 0 (constant output)
- **Model 1 (Linear) - Though we may expect non-linear data, often at 6,000 hours the curve has not yet formed making this often a good fit model**
- **Model 2 (Simple Exponential)**
- **Model 3 (Linear + Exponential) * - Asymptotes and often gives infinite L_{70} predictions**
- **Model 4 (Log) * - L_{70} results often unreasonable (L_{70} is estimated to be 100s of years)**
- **Model 5 (Linear + Log) - Sensitive to data noise which leads to increased uncertainty in L70 predictions**
- Model 6 (Exponential + Log) - Fits to the warm-up hump and can "artificially" model accelerated or no decay (infinite L_{70})
- Model 7 (Linear + Exponential + Log) - Fits to the warm-up hump and can "artificially" model accelerated or no decay (infinite L_{70})

* Included as candidate models, but will generally not be selected.

Pacific Northwest NATIONAL LABORATORY

Model Evaluation Routine & Results

Based on initial evaluation of model characteristics, Models 1-5 were chosen for further detailed evaluation

Evaluation Routine:

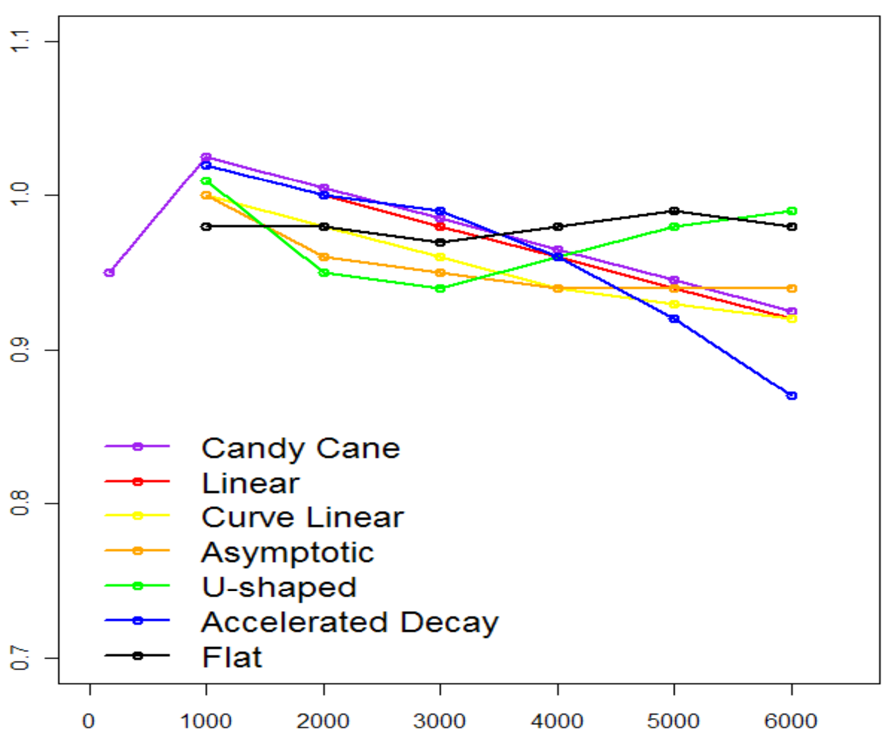
1. Sample or measured LED data is obtained
2. Uncertainty in measurements ("noise") is identified (1% noise default).
3. 500 different potential LED decay observations are simulated representing noise parameters.
4. All 5 models are fit to each simulation of LED decay observations.
5. SSE and L_{70} prediction values are stored for each simulation.
6. Evaluations are made on the set of simulations.

Pacific Northwest NATIONAL LABORATORY

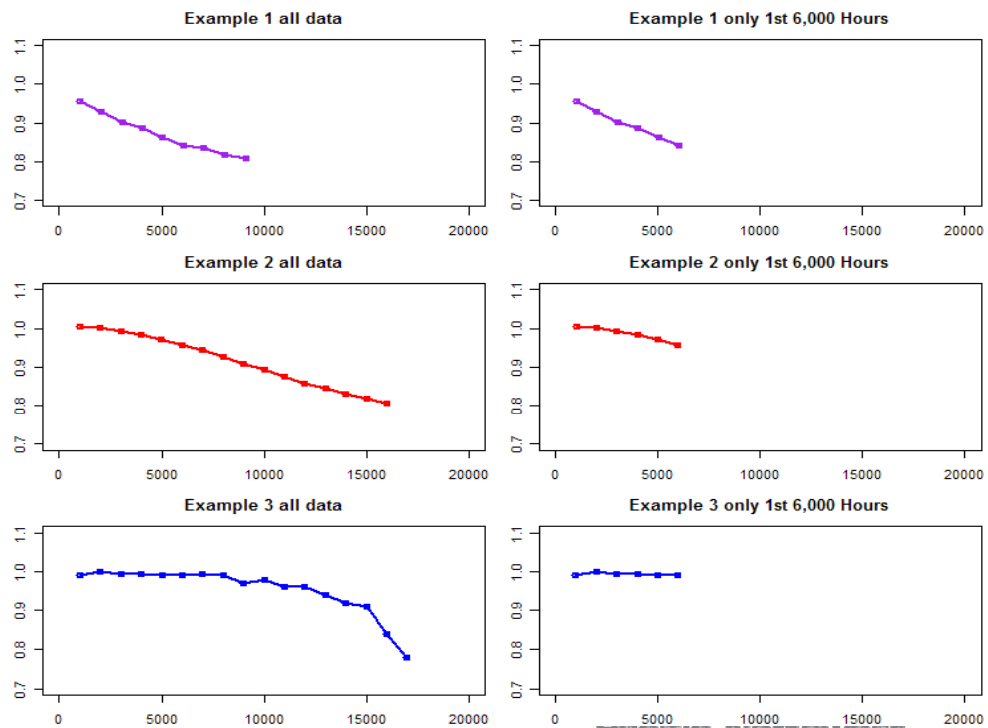
Evaluation Routine & Results based on Subset Data

- Evaluate process and results
  1. Using “Toy” examples (every 1,000 hours up to 6,000)
  2. Comparing long real data sets to their 6,000 hour (LM-80) data subsets

1

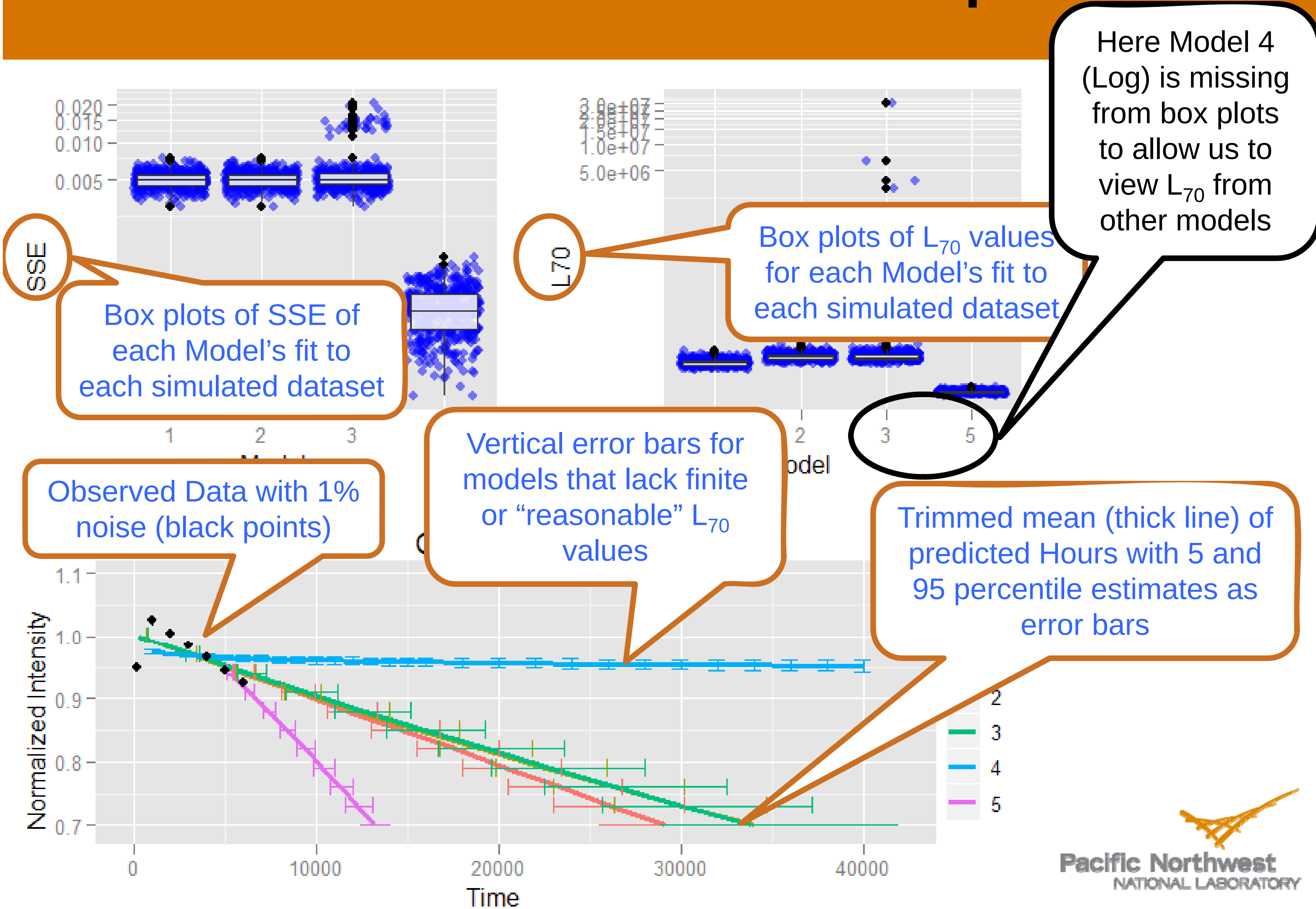


2



Pacific Northwest
NATIONAL LABORATORY

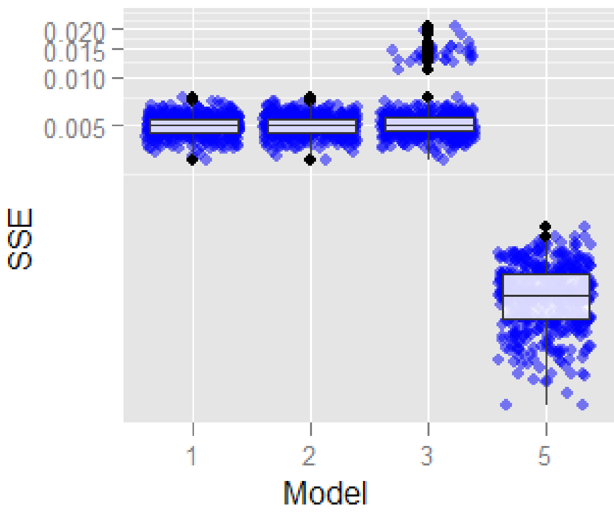
Evaluation Results Description

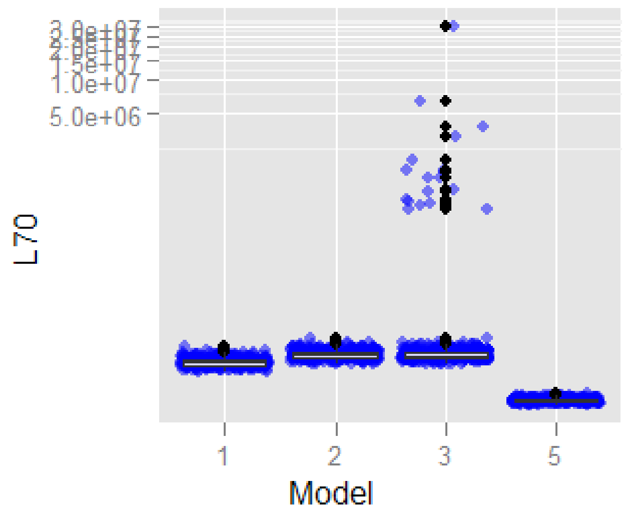


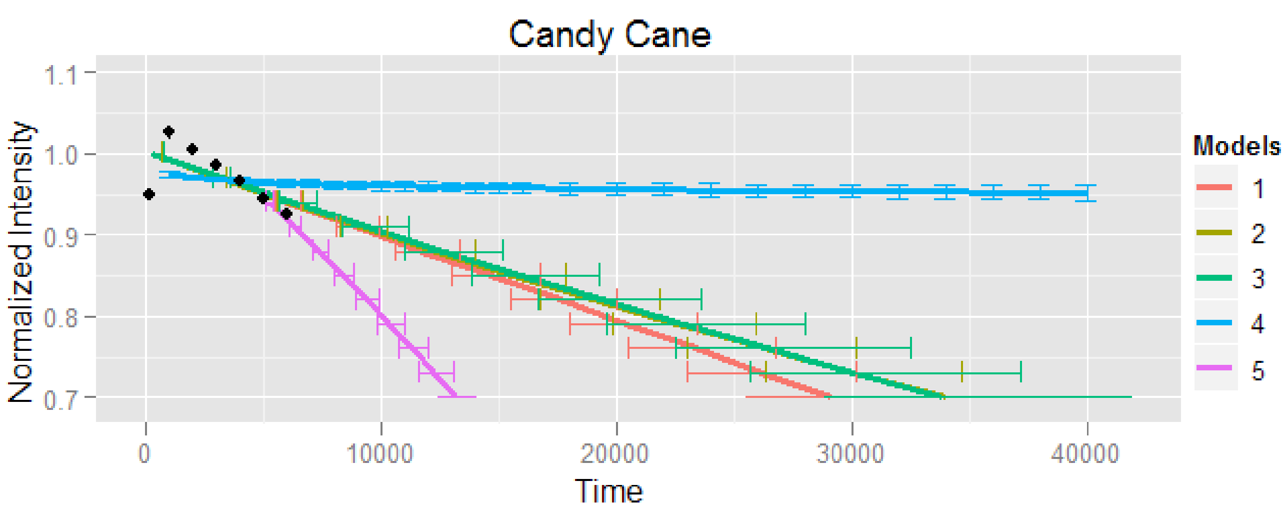
Evaluation Results – Sample Data

(every 1,000 hours up to 6,000)

Toy Examples- "Candy Cane"







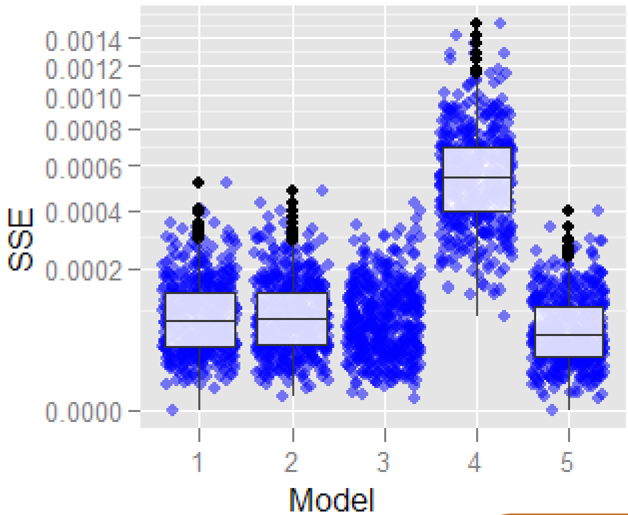
- Model 5 (linear-log) will fit the best and predict the lowest L_{70} values.
- All other models will predict smaller L_{70} values when the first point is excluded.
- Early measurements before the LED has "warmed-up" should not be included in the modeling.

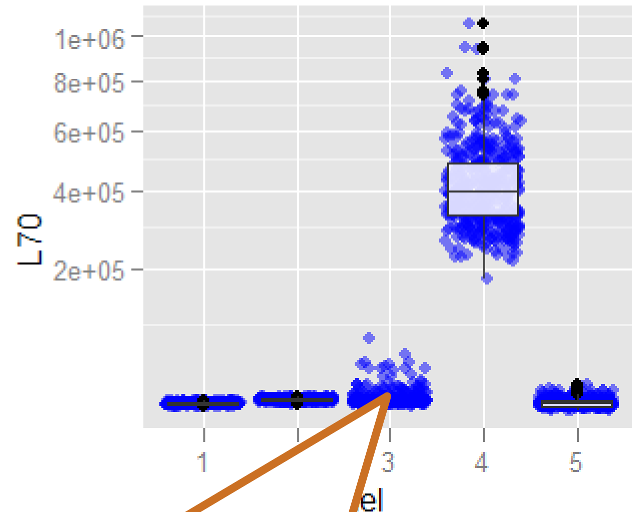
Pacific Northwest NATIONAL LABORATORY

Evaluation Results – Sample Data

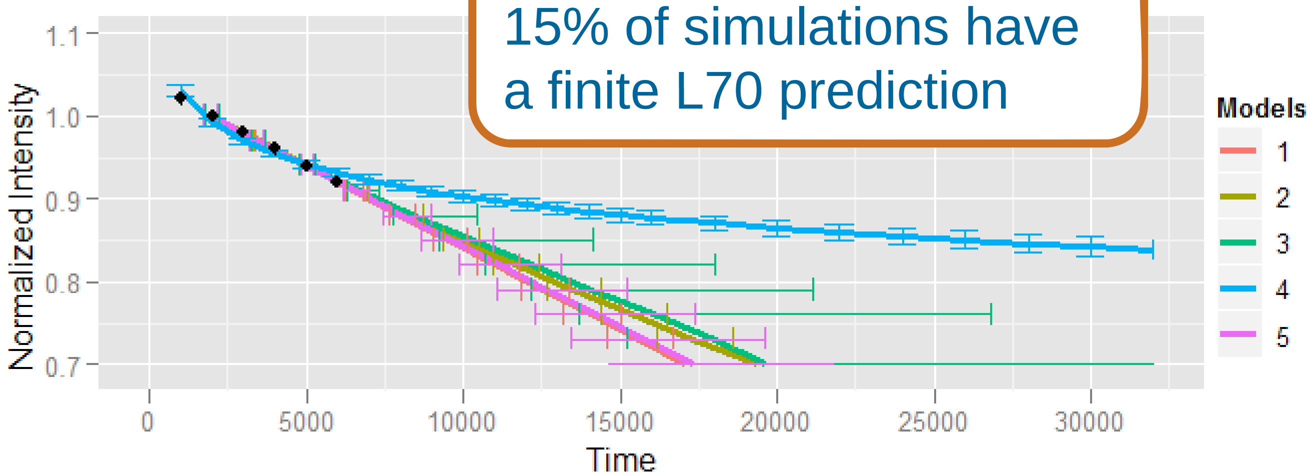
(every 1,000 hours up to 6,000)

Toy Examples- “Linear Data”





No box drawn if less than 15% of simulations have a finite L70 prediction



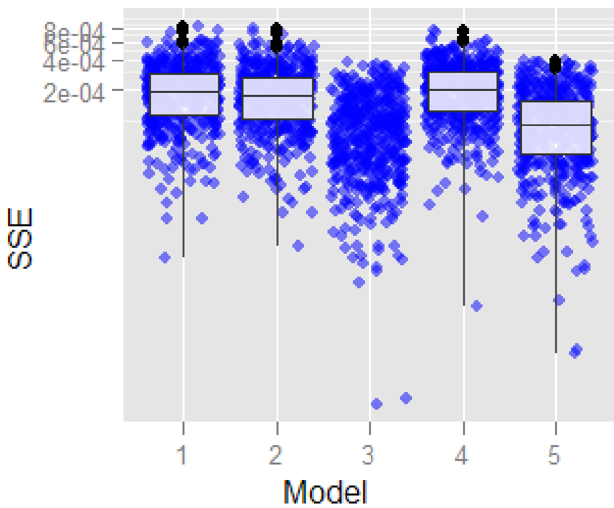
- Most models will generally predict L_{70} values in the same region – Model 4 (log) is the exception.
- Model 1 (linear) and 5 (linear-log) will predict very similar values.

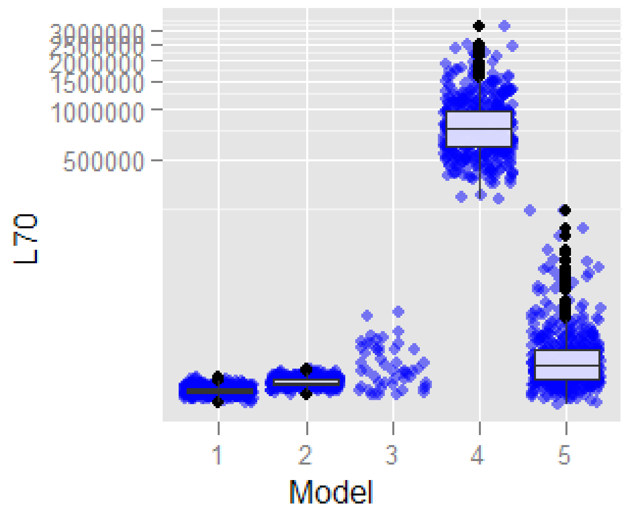
Pacific Northwest
NATIONAL LABORATORY

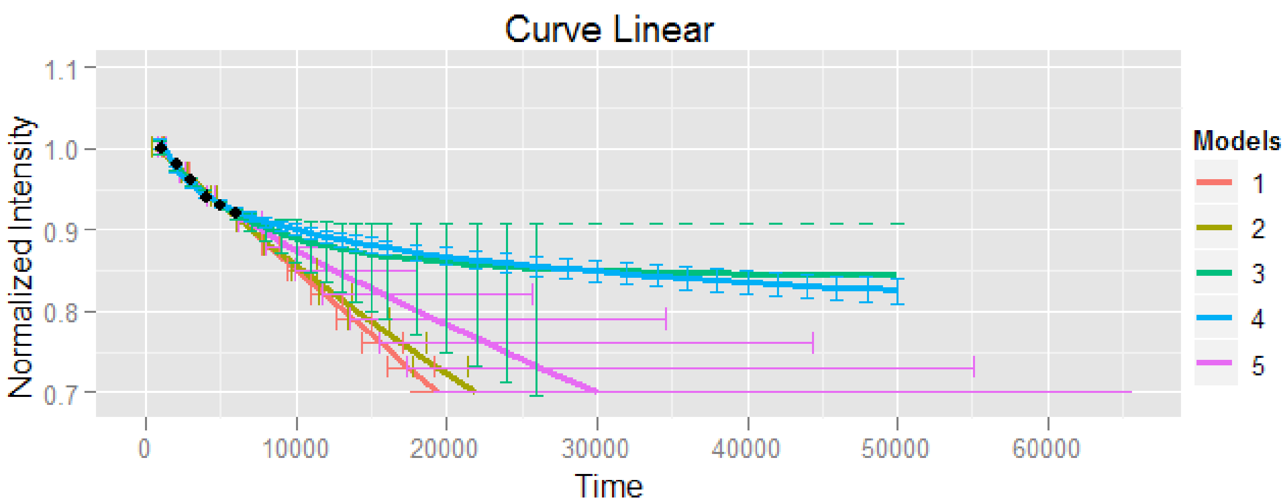
Evaluation Results – Sample Data

(every 1,000 hours up to 6,000)

Toy Examples- “Curve Linear”





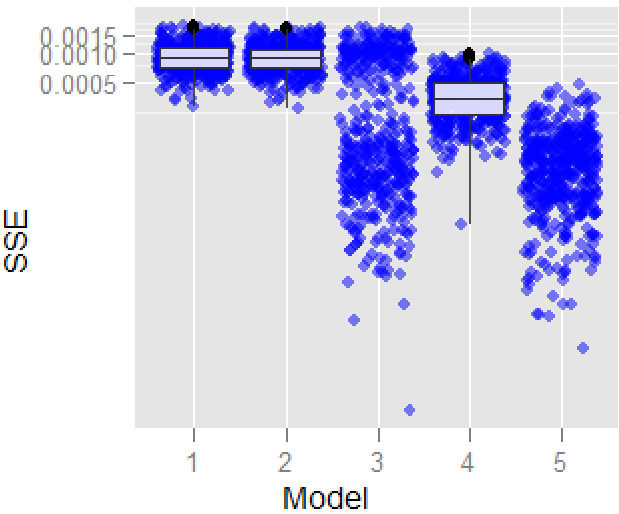


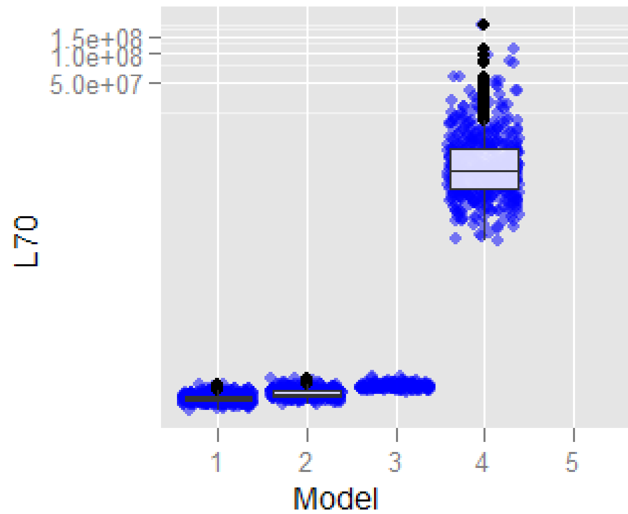
- All models fit the data roughly the same (similar SSE)
- Model 3 (linear-exp) tends to asymptote above L_{70} .
- A wide range of predicted L_{70} values.

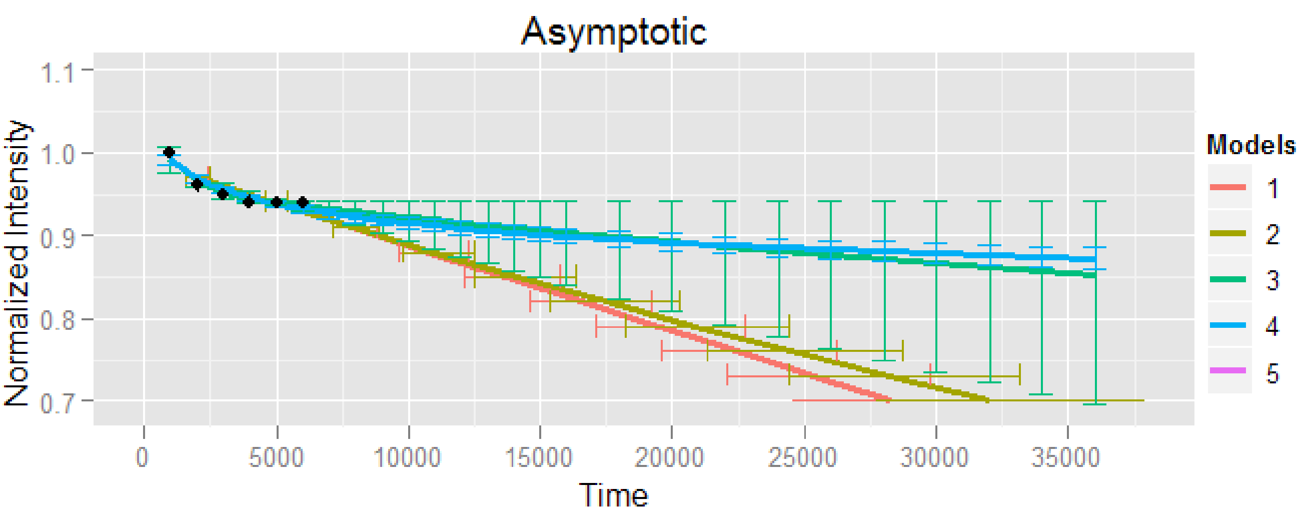
Evaluation Results – Sample Data

(every 1,000 hours up to 6,000)

Toy Examples- "Asymptotic"





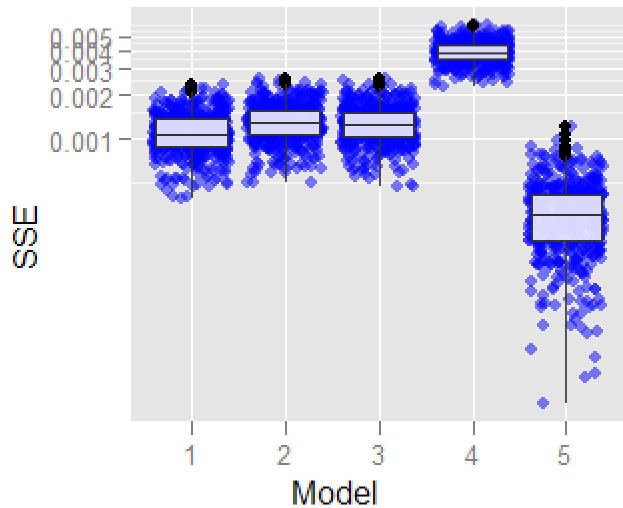


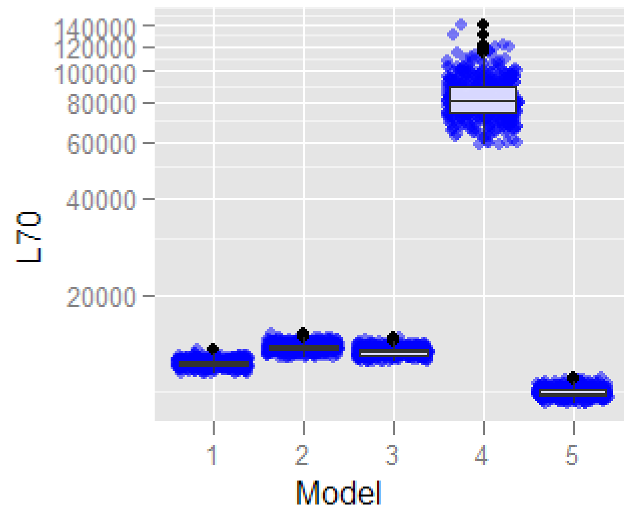
- Model 3 (linear-exp) either asymptotes above L_{70} or is very similar to Model 1 (linear) (see two clusters in SSE plot).
- Model 4 (log) estimates L_{70} over 1,000,000 hours
- Model 5 (linear-log) generally asymptotes above L_{70} (not shown).

Evaluation Results – Sample Data

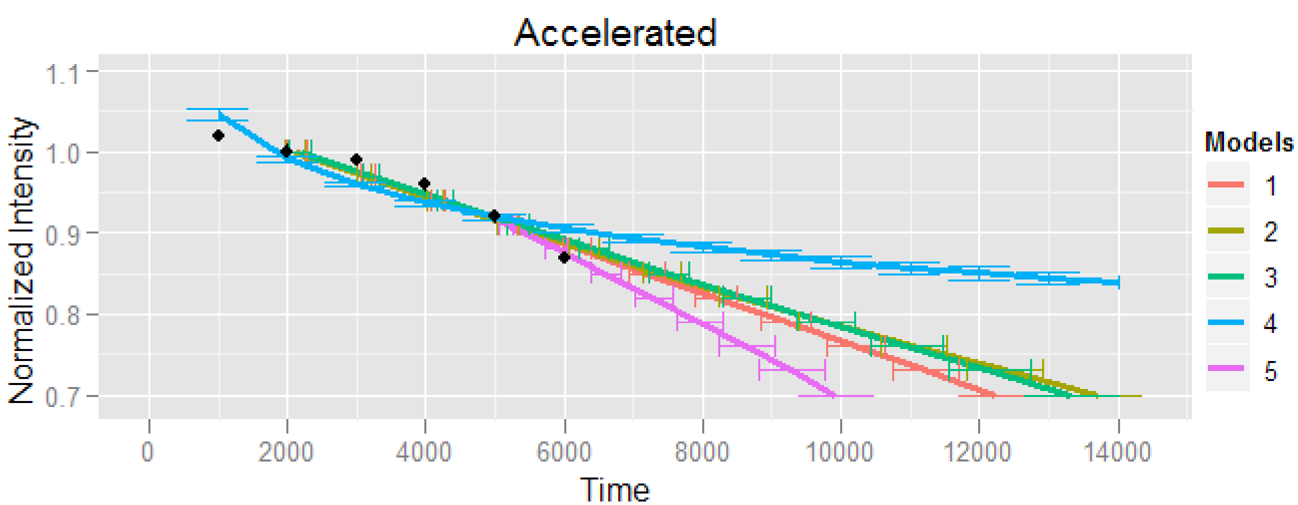
(every 1,000 hours up to 6,000)

Toy Examples- "Accelerated"





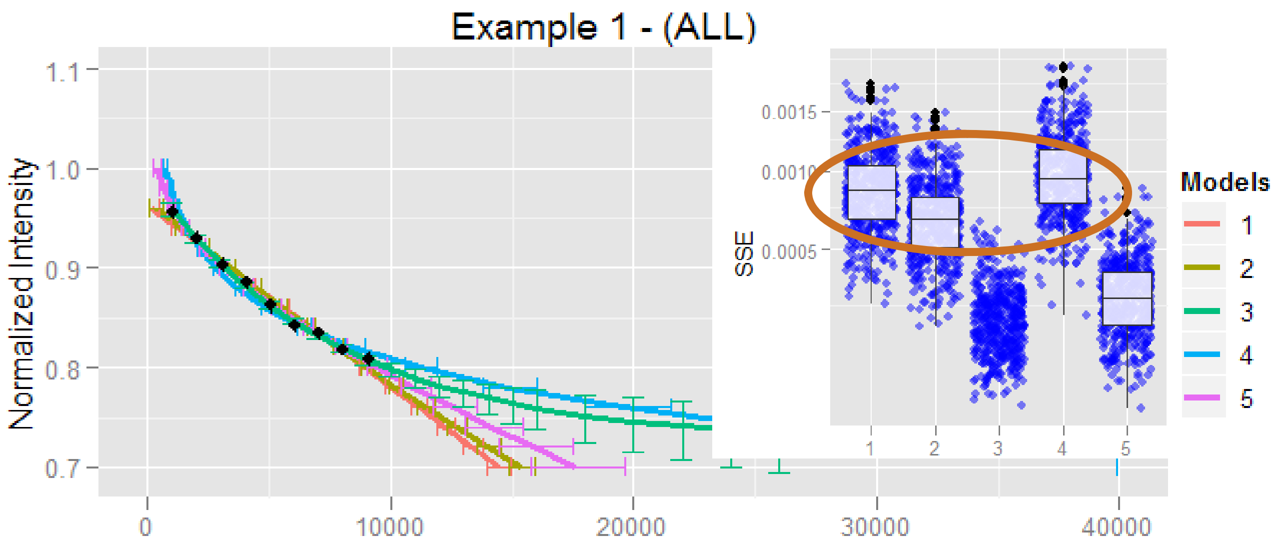
- Model 5 (linear-log) will best fit accelerated decay
- Model 5 generally gives the lowest predicted L_{70} values
- Model 3 (linear-exp) asymptotes.

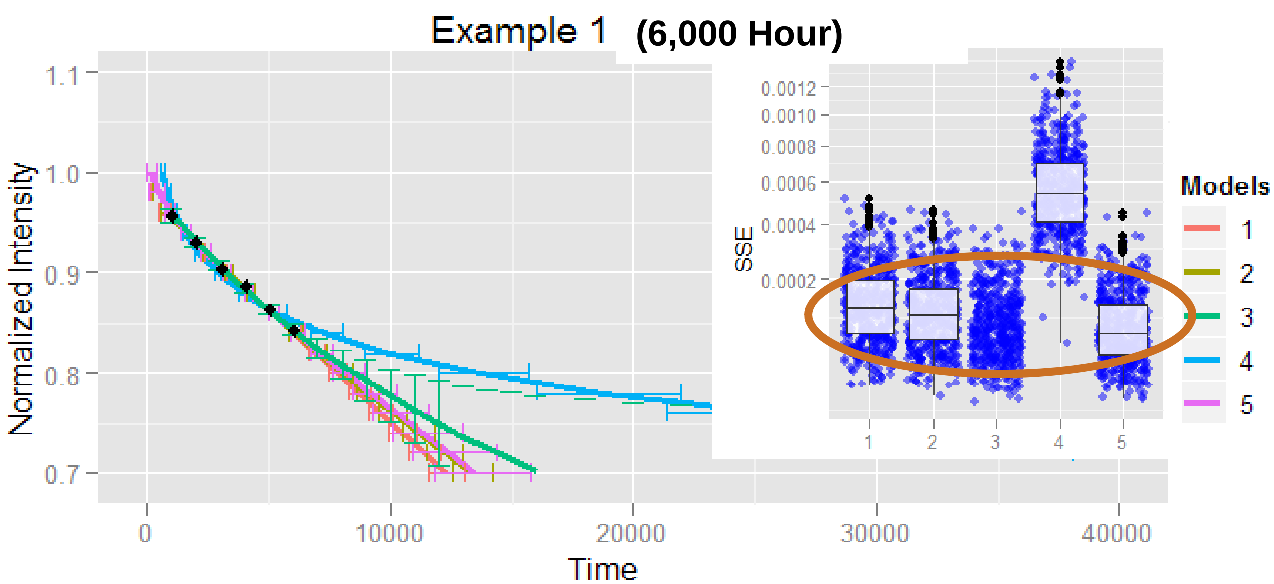


Pacific Northwest NATIONAL LABORATORY

Evaluation Results – Real Data Full and Subset Comparison

Full Data Set vs. 6,000 Hour Subset

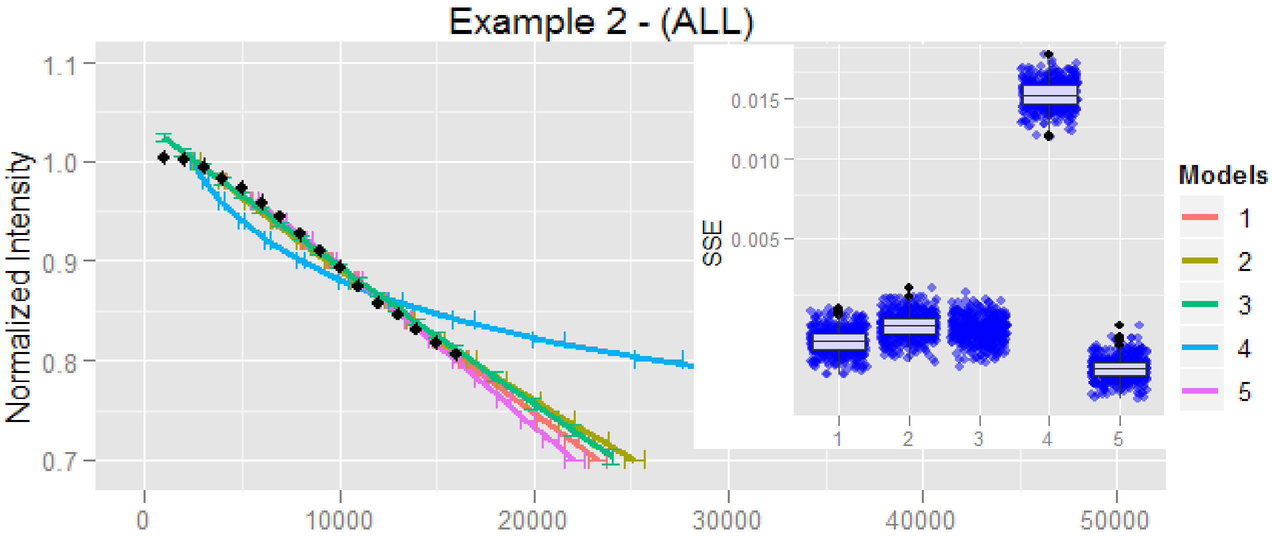


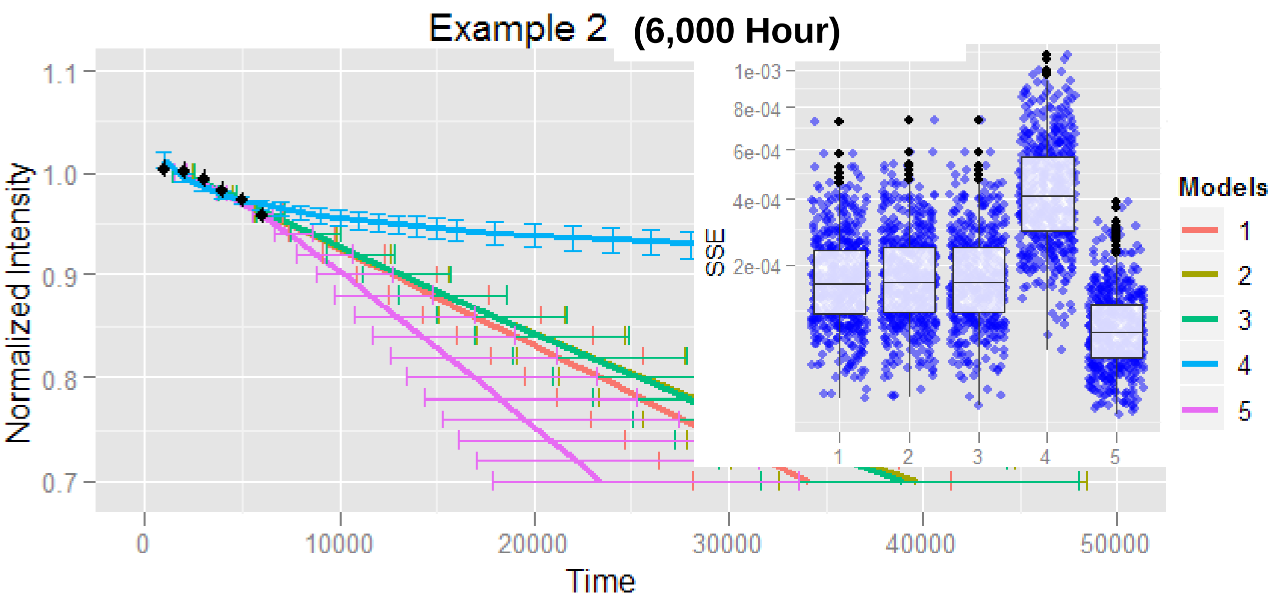


- With additional data (3 more observations out to 9,000 hours) Models 1 (linear) and 2 (exponential) are no longer considered as in the group of the best models.

Evaluation Results – Real Data Full and Subset Comparison

Full Data Set vs. 6,000 Hour Subset

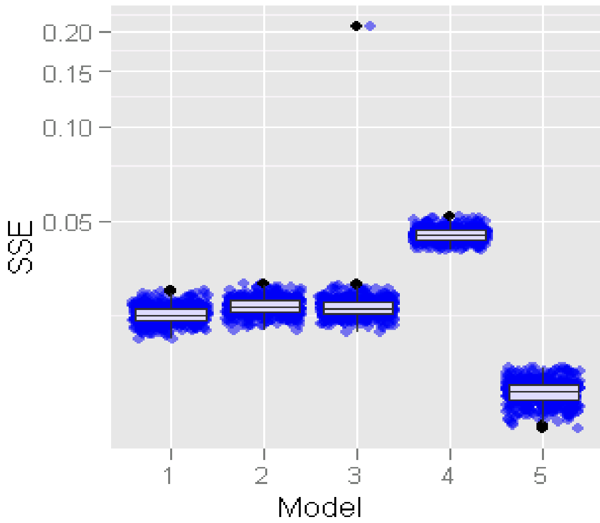


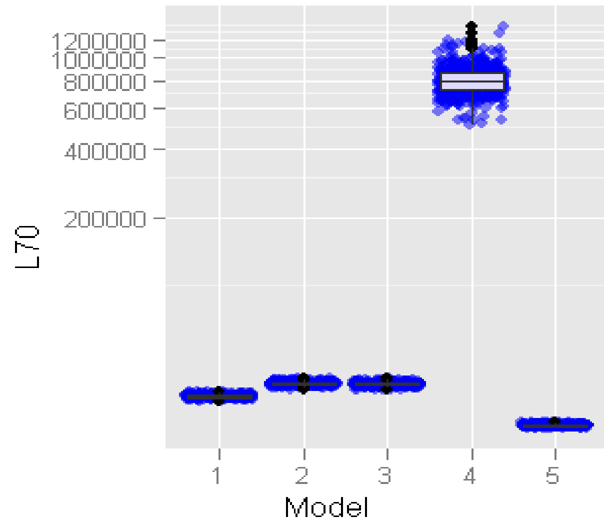


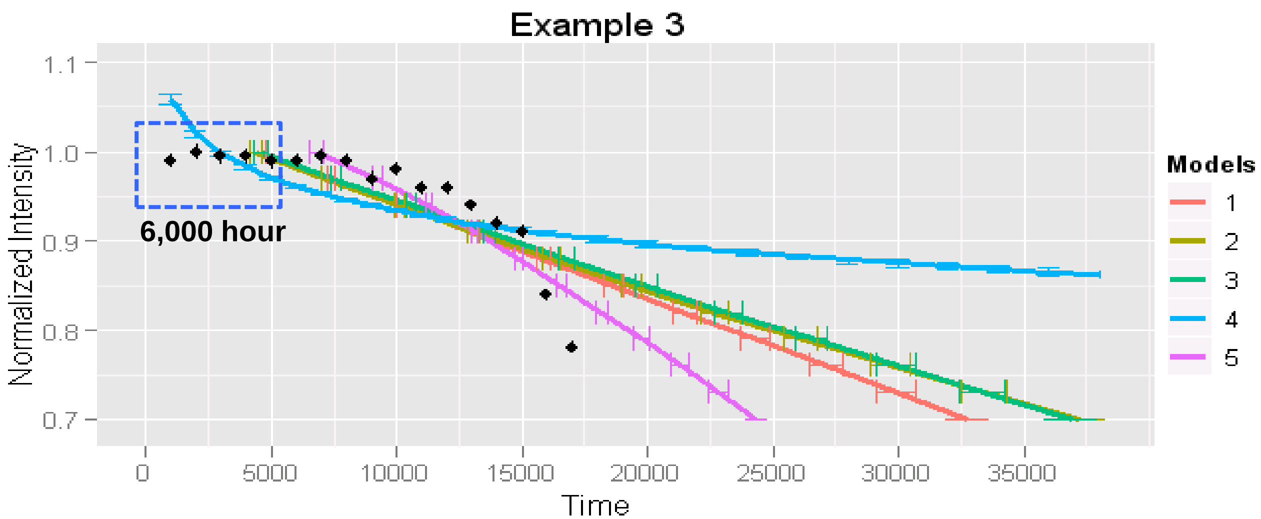
- Model 5 (linear-log) could be argued to have the minimum SSE (implying best fit) for both full and 6,000 hour subset.
- Average predictions of Model 5 for Full Data and 6,000 hour subset are similar.
- **…But uncertainty is reduced in the L_{70} predictions when using the Full Data Set.**

Evaluation Results – Real Data Full and Subset Comparison

Full Data Set vs. 6,000 Hour Subset

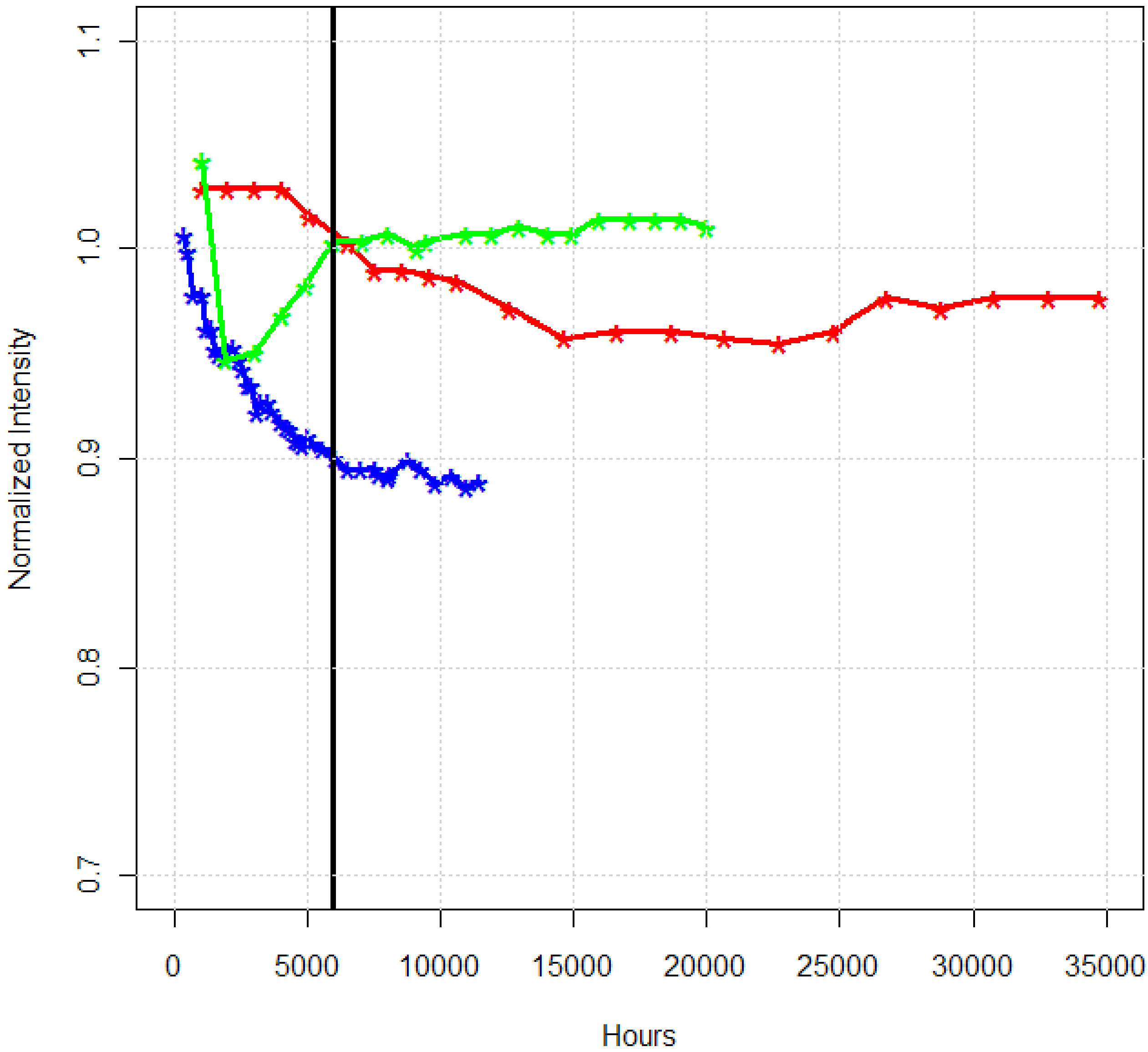






- All five models fit poorly to the Full Data Set.
- 6,000 hour subset is essentially a flat line, to which none of the models will provide a reasonable L_{70} estimate.
- Model 5 (linear-log) shows again how it tends to fit accelerated decay structures “best”.

Additional Full vs. Partial Dataset Examples



- Often don't see the curvature by 6,000 hours (black vertical line)
- Red would likely predict an accelerated decay
- Green can't be fit using the models (would give infinite L_{70})
- Blue shows curvature, but can't see the decelerated decay yet

Pacific Northwest
NATIONAL LABORATORY

Conclusion: Observations using LM-80 Type Data (every 1,000 hours up to 6,000)

- Models 6 & 7 are not expected to provide useful results for LM-80 data and have not been further evaluated.
- Model 4 often provides infinite L_{70} predictions.
- Model 3 & 4 often asymptote.
- Most models behave poorly if LED “warm-up” (data rising prior to descending decay) measurements are included.
- Extending the time period of observations will provide improved predictions of the L_{70}
  - Reduction in uncertainty by fitting more data points, or
  - True decay structures start to manifest
- **Suggest**: Choose the best model and L_{70} projection based on a conservative set of “rules”.

Proposed L70 Projection "Rule" Set

1. If there is a maximum lumen output point that exceeds the initial lumen output by more than the data uncertainty (**1**% for example), then remove the "warm-up" measurements before the maximum lumen output point.
2. Fit models 1, 2, 3, 4, & 5 to remaining data and simulate **500** times with rated laboratory uncertainty (propose **1**% as a default).
3. Calculate SSE and L_{70} values of **500** simulations for each model.
4. If the SSE of one model is smaller than those of the other models without overlapping them by more than **10**%, use the L_{70} from that model (technically a statistical comparison of SSE).
5. If the SSE values of several models are smaller than those of the other models without overlapping them more than **10**%, but there is no significant separation among this group, use the smallest L_{70} value of all models in this group.
6. If the SSE values of all models overlap more than **10**%, use the smallest L_{70} value of all models.
7. Use the lower **25th** percentile of the selected L_{70} values (based on steps 4, 5, 6 above) as the L_{70} projection for this set of LED data.
8. Apply **6** times rule. If the L_{70} projection is > 6 times the measurement period, use 6 times the measurement period as the final L_{70} projection.

Pacific Northwest NATIONAL LABORATORY

- Questions?