Applying Information-Theoretic Approaches for Objective Model Selection and Quantification of Model Selection Uncertainty

Odele Coddington[1], Peter Pilewskie[1,2], Erik Richard[1], and Andrew Kren[1,2]
(1) Laboratory for Atmospheric and Space Physics (LASP), (2) Department of Atmospheric and Oceanic Science (ATOC): CU-Boulder
Contact Info: odele.coddington@lasp.colorado.edu or by phone (303) 492-9318

Laboratory for Atmospheric and Space Physics
University of Colorado **Boulder**

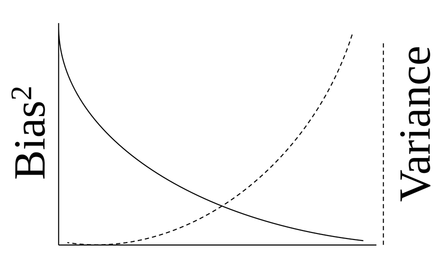
OVERVIEW

What Makes a Model the "Best Model"?

Valid inferences from scientific data may depend on a model of information in the data. There are three aspects to the general problem of valid inference [*Fisher*, 1922].

1. Model Specification*	2. Model Parameter Estimation	3. Estimation of Precision

***Preceded by formulation of a set of candidate models guided by scientific experience, debate, literature, and experimentation. Good data and good candidate models, where each model represents a scientific hypothesis, can provide insight into the physics of an underlying problem.



A good fit is necessary but not sufficient! The "best" model will be near the intersection of these curves.

The Principle of Parsimony – A conceptual trade off between squared bias and variance versus the number of model parameters. A component of all model selection approaches.

Information-Theoretic Methods of Model Selection

Kullback-Leibler Distance [*Kullback and Liebler*, 1951]- The negative entropy (i.e. relative entropy); a measure of (information) distance between two models. Derived from C. Shannon [1949] *Theory of Communication* and Boltzmann's general theory of thermodynamic entropy.

$$S = \sum p(x_i) \ln \frac{p(x_i)}{q(x_i)}$$

K-L Distance always positive, unless there is a perfect match between model 'q' and truth 'p' (in which case, K-L distance = 0).

An Information Criterion (AIC) [*H. Akaike*, 1973]

Links Maximum Likelihood Estimation (£) and Probability Theory with K-L Distance.
Useful for statistical inference of time-series data such as gap-filling, combining data sets, and multimodel averaging and quantification of model selection uncertainty

AIC = -2 log(£) + 2K, or
AIC = -2 log(£) + 2K * [N/(N-K-1)] for small sample size [Suguira, 1978]

K = No. of model parameters, N=Sample Size

Description of AIC "Body Of Evidence" Metrics

Metric Name	Derivation	Information-Theoretic Meaning
AIC_{min}	Minimum of AIC values	The "best-fit" model.
Δ AIC	AIC - AIC_{min}	Level of support (0-2: substantial; 4-7: considerably less; > 10: essentially none
Model Likelihood	Proportional to exp(-0.5*Δ AIC)	Relative strength of evidence for a model
AIC Weights	Model Likelihood Normalized by Total Likelihood	Improves interpretation of model likelihood; an effective way to scale Δ AIC values for multimodel averaging and quantifying model selection uncertainty
Evidence Ratio	AIC Weights normalized by AIC Weight (of AIC_{min} model)	Weight of evidence with respect to the best model

ON-ORBIT DEGRADATION OF SORCE SIM

Space optics exposed to harsh solar radiation degrade as a function of wavelength and exposure amount. Time-series of spectral irradiance must be corrected for this degradation to achieve record stability. Correction methods require the assumption of a degradation model.

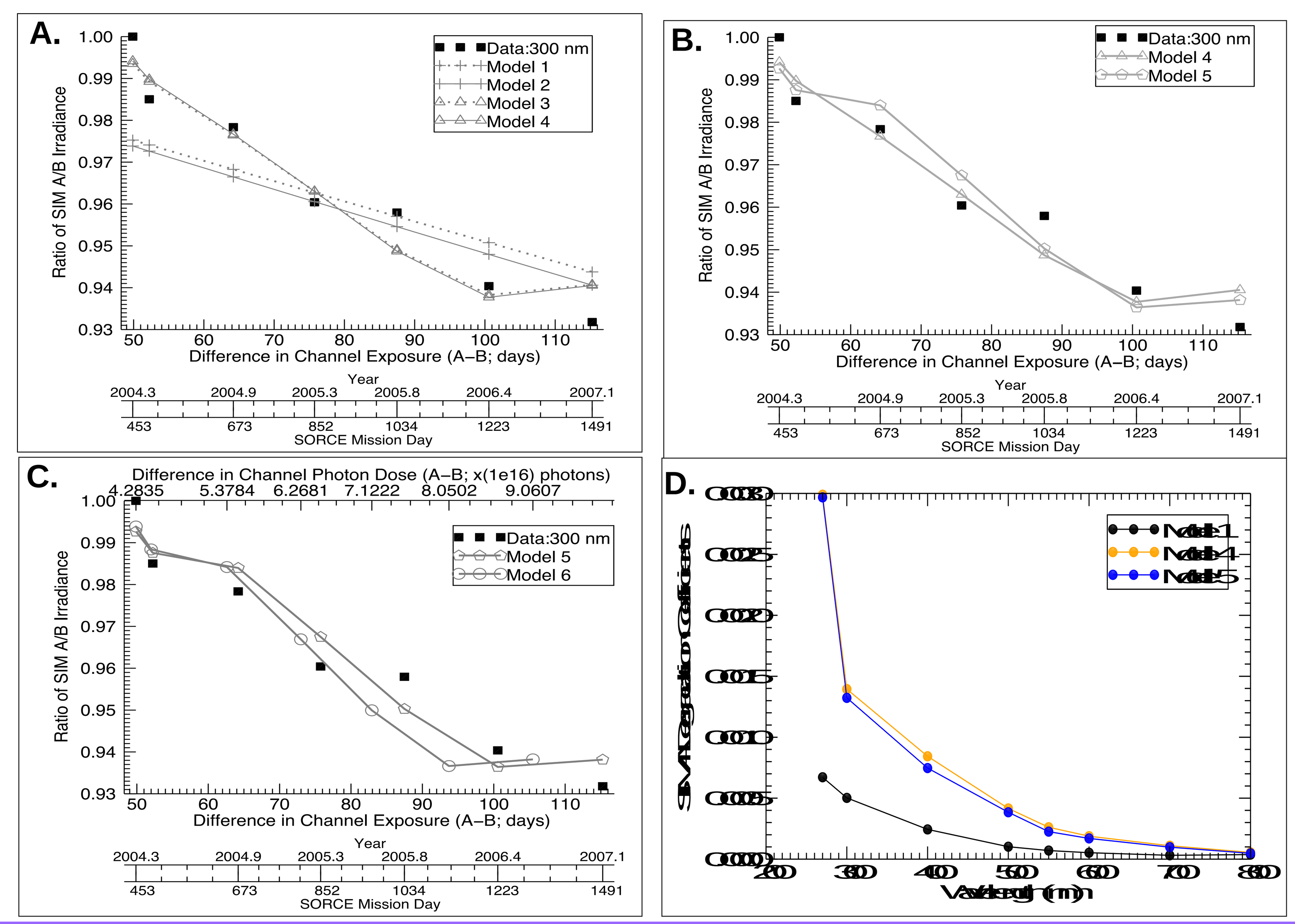
Motivation

The sources of degradation are likely complex, with many small effects and interactions, that will (likely) remain unknown to us.

Science Question

What insight into the physical processes underlying SIM on-orbit degradation can be gained through a ranking and scaling model "usefulness" using An Information Criterion (AIC)?

RESULTS & INTERPRETATION



1. Incomplete removal of solar variability impacts the best-fit degradation function (fig A.; Model #1-2 curves), but little impact when degradation function assumed to vary between channels (fig A.; Model #3-4 curves).
2. Uncertainty in solar exposure (independent variable) results in bias in degradation function (fig B., fig D.; Model #4-5 curves).
3. Using photon "dose" (Model #6) can hasten or decrease degradation relative to exposure "time" (Model #5), depending on solar cycle activity and proximity to Sun (fig C.).
4. Optical degradation is a smooth function of wavelength (fig D.), with greatest degradation at shortest wavelengths and near 0 longward of 800 nm.

Model (#K)	AIC Value	Δ AIC	Likelihood	AIC Weights	Evidence Ratio	R^2 (%)	Adjusted R^2 (%)
1 (3)	70.431	5.500	0.064	0.060	15.642	69.6	63.5
2 (3)	64.931	0.0	1.000	0.940	1.000	68.6	62.3
3 (4)	87.840	22.909	0.0	0.0	9.4e4	93.5	90.2
4 (4)	82.831	17.900	0.0	0.0	7.7e3	93.4	90.2
5 (5)	125.025	60.09	0.0	0.0	1.1e13	93.8	85.7

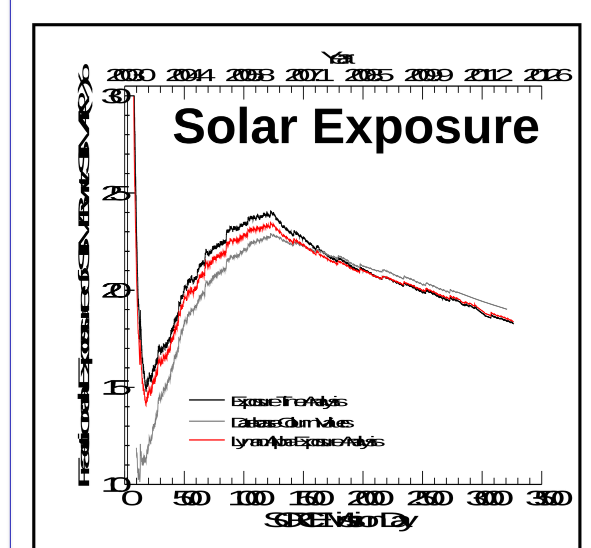
1. Model #2 is the AIC "Best" model at all wavelengths.
2. Gains made in best-fit for models #3-#6 are (more) offset by penalties from increased number of fit parameters.
3. Model #2 has the worst "fit" (R^2 value) at all wavelengths, highlighting the importance of model selection vs. model description.
4. Results emphasize the importance of sample size; increasing sample size critical for using information-theoretic approaches for valid inference.
5. Future work:
-Expand sample size with table scan data (currently used for degradation monitoring) and/or diode data.
-Use AIC weights to derive multimodel averages and provide a model selection uncertainty to corrected time-series estimates of precision.

ALGORITHM INPUTS

Data, Knowledge of Uncertainties, and Solar Exposure

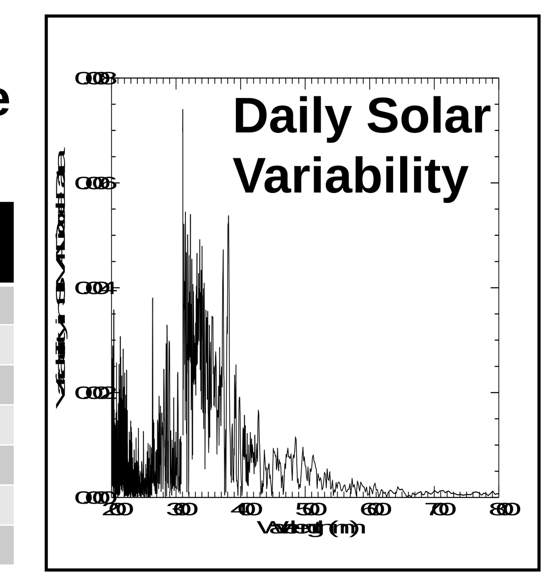
- SIM Ver. 19 data with 1-AU and temp. corrections applied.
- Binned irradiances (270, 300, 350, 400, 500, 550, 600, 700, and 800 nm; variable bin size) obtained from ESR Full Scans measured by A/B channels on same day ("golden" days 453, 491, 673, 852, 1034, 1223, and 1419).
- Measurement uncertainty = 0.2% (1-σ); assumed constant with time and wavelength.
- Uncertainty d/t incomplete removal of solar variability determined from ratio'd irradiance– variable with time and wavelength; obtained from 15 orbits of diode measurements on golden days.
- Total Measurement Uncertainty = $\sqrt{(\sigma^2_{meas} + \sigma^2_{sun})}$
- Solar Exposure time (days) obtained through counting seconds where i) shutter open, ii) pointing on Sun, and iii) hard-radiation trap "out".
- Solar Exposure uncertainty estimated (upper limit) from cumulative time of missing science and instrument housekeeping packets.
- Measure of photon "dose" obtained using SORCE Lyman-alpha time series (no 1-AU correction) and knowledge of solar exposure time and SIM A/B entrance slit area.

Coddington et al., 2012, A new look at solar exposure and SORCE degradation, SSI Trends Workshop II, Annapolis, MD, Sept. 2012



Solar Exposure Uncertainty

Mission Day	SIM A (Days)	SIM B (Days)
453	0.78	1.04
491	0.78	1.04
673	2.41	2.83
852	2.44	2.87
1034	2.47	2.90
1223	2.49	2.91
1419	2.66	3.08



HYPOTHESIS-BASED APPROACH TO MODEL SPECIFICATION

Underlying Premise(s) to Degradation Monitoring -
•The irradiance in each SIM channel (A/B) is a function of time, wavelength, and the optical degradation, which is dependent upon exposure (the rates of exposure differ by channel).
•Instrumental degradation function is exponential in form.
•By analyzing ratios of A/B irradiance, we remove the solar variability from the measured signal.

$$I_A / I_B \propto \frac{S(t)/\deg_A(t^A_{exp})}{S(t)/\deg_B(t^B_{exp})}$$

The ratio of SIM A/B Irradiance is proportional to the ratio of the instrument degradation.

Candidate Degradation Models

Model #1 – The degradation function is the same in both channels.
Model #2 – Same degradation function but solar variability not removed.
Model #3 – The degradation function is different between channels and solar variability is removed.
Model #4 – Channel-dependent degradation functions and solar variability is not removed.
Model #5 – Channel-dependent degradation functions, solar variability not removed, and exposure time is uncertain (and uncorrelated to measurement uncertainty).
Model #6 – Same as Model #5, but solar exposure is measured by (uncertain) photon "dose".