A predictor analysis framework for surface radiation budget reprocessing using satellite data

Patricia A. Quigley*

Science Systems and Applications, Inc. (SSAI),
Hampton, VA 23666, USA
Email: patriciaq222@gmail.com
Email: patriciaq@gmail.com
*Corresponding author

Resit Unal

Old Dominion University (ODU),
Norfolk, VA 23529, USA
Email: runal@odu.edu

Paul W. Stackhouse Junior

National Aeronautics and Space Administration (NASA),
Hampton, VA 23666, USA
Email: paul.w.stackhouse@nasa.gov

Stephen J. Cox

Science Systems and Applications, Inc. (SSAI),
Hampton, VA 23666, USA
Email: stephen.j.cox@nasa.gov

Abstract: Equipped with various types of imagers, lasers and radars, dozens of satellites orbit the earth every day collecting and relaying data for weather and atmospheric analysis, communication and navigation applications and planetary studies. Earth orbiting satellites are part of the critical space infrastructures. NASA's Global Energy and Water Cycle (GEWEX) surface radiation budget (SRB) shortwave algorithm derives long-term datasets from satellite data of the distribution of the sun's energy to the surface and back to space. This paper presents an analysis framework to describe propagation of input parameter variability to output data results in algorithmic computations, and then quantify the variability in the solution sets. The SRB shortwave algorithm and design of experiments (DOE) methods are utilised to determine significant input parameters and interactions. A sensitivity analysis is also conducted to determine the variability in the output data for each dependent variable varying within their range using Monte Carlo simulation.

Keywords: surface radiation budget; SRB; variability; design of experiments; DOE; augmented minimum point designs; GEWEX SRB.

Reference to this paper should be made as follows: Quigley, P.A., Unal, R., Stackhouse Jr., P.W. and Cox, S.J. (2021) 'A predictor analysis framework for surface radiation budget reprocessing using satellite data', *Int. J. Critical Infrastructures*, Vol. 17, No. 1, pp.71–85.

Biographical notes: Patricia A. Quigley completed her PhD at Old Dominion University. Her main interests involve robust engineering design solutions, response surface, systems analysis, process optimisation, multi-criteria decision analysis and risk analysis. She is currently a functional test engineer contractor at the Atmospheric Science Data Center (ASDC) at Langley Research Center (LaRC) in Hampton, Virginia, and is coding software in support of the SRB global analysis.

Resit Unal is a Professor in the Department of Engineering Management and Systems Engineering at Old Dominion University (ODU) in Norfolk, Virginia, USA. He is the past Chair of the same department at ODU and Fellow of the American Society for Engineering Management (ASEM). He holds a BS in Electrical Engineering from Middle East Technical University (METU), Ankara, Turkey and received his MS and PhD in Engineering Management from the University of Missouri-Rolla, Rolla, Missouri, USA. His research interest includes multidisciplinary design optimisation, robust design, quality engineering, response surface methods, and parametric cost estimating. His research has been published in the *Engineering Management Journal*, *IEEE Transactions in Engineering Management*, *Quality Engineering*, *Journal of Cost Analysis and Management*, *Reliability Engineering & System Safety*, *Journal of Parametrics* and *Journal of Spacecraft and Rockets*.

Paul W. Stackhouse Junior is a Senior Research Scientist at the NASA Langley Research Center in Hampton, VA. His research interests encompass the derivation and assessment of global solar and thermal infrared radiative flux estimates from satellite measurements and analysis. He also leads projects aiming to provide datasets from this research to those working in energy and agricultural applications.

Stephen J. Cox is a Lead Research Scientist at SSAI, Inc. in Hampton Virginia. He works primarily on the NASA/GEWEX Surface Radiation Budget project, which derives a global history of solar and thermal infrared radiative fluxes from satellite measurements. His research interests include solar atmospheric radiative transfer, and cloud and aerosol radiative effects.

This paper is a revised and expanded version of a paper entitled 'Identifying key variables on surface radiation budget using design of experiments methods' presented at the 39th American Society for Engineering Management Conference, Coeur D'Alene, Idaho, 17–20 October 2018.

1 Introduction

The Earth's radiation budget is an accounting of all incoming energy from the sun and outgoing energy reflected and radiated to space by earth's surface and atmosphere (Barkstrom and Smith, 1986). The National Aeronautics and Space Administration (NASA)/Global Energy and Water Cycle Exchange (GEWEX) surface radiation budget (SRB) project produces and archives long-term datasets representative of this energy

exchange system on a global scale (Hinkelman et al., 2009; Stackhouse et al., 2011). The goal of this project is to produce reliable, globally derived top-of-atmosphere (TOA), within atmospheric and surface radiative flux products relevant to the exchange of radiative energy between space and the Earth. Properly calibrated ground sites measurements can obtain direct measurements of the downward fluxes of energy to the surface of the Earth accurately. However, since it is not practical to cover the entire globe, including the oceans, with ground data measurement sites, it is desirable to use satellite analysis to estimate this information to obtain estimates of the global data quantities (Whitlock et al., 1994). Satellite instruments also provide observations that are used to directly estimate the TOA information which ground sites cannot. The knowledge of atmospheric properties and their fluxes is beneficial to many areas of research, including determining the overall radiation budget of the planet and it's variability (Stephens et al., 2012), the design of renewable energy systems and the provision of solar energy information for crop modelling (NASA, 2017a, 2017b, 2017c) having sustained benefits to industries from shipping to farming.

2 Problem statement

Satellite coverage for data collection overcomes the problems associated with ground site positioning, providing an opportunity for obtaining a global estimate. However, surface irradiance cannot be measured directly from the TOA by satellites. Therefore, it must be derived from a variety of satellite acquired atmospheric data using several algorithms. For this purpose, the SRB algorithm was developed (Whitlock et al., 1994). The SRB shortwave algorithm uses radiative transfer algorithms to derive radiative fluxes from satellite measurements and associated atmospheric and surface data products estimating the presence and optical thickness of clouds, aerosols and atmospheric gaseous constituents (Stackhouse et al., 2011). The SRB product is comprised of the shortwave (wavelengths from 0.2 to 4.0 μm) and longwave (4.0μm–∞) radiative components of the system. Here we examine only the shortwave data. It is stored as 3-hourly, daily, monthly/3-hourly, and monthly averages of 1° × 1° grid cells (EOSWEB, 2017).

Research indicates that variability in the input data used by the algorithms is a key source of uncertainty in the resulting output data (Hinkelman et al., 2009; Zhang et al., 2007, 2013, 2015; Rossow and Zhang, 1995; Stephens et al., 2012; Raschke et al., 2016). As a result, sensitivity studies have been conducted to estimate the effects this variability has on the output datasets using linear techniques (Kato et al., 2012). This usually entails varying one input parameter at a time while keeping all others constant or by increasing all input parameters by equal random percentages, in effect changing input values for every cell for every three hour period and for every day in each month. This approach can equate to millions of independent changes in input without taking into consideration the interdependences or interactions among the input parameters.

To overcome this challenge, a more comprehensive approach was adopted in this study for evaluating the SRB algorithm to identify both the input parameters and parameter interactions that most significantly affect the output data for global estimates. Design of experiments (DOE) techniques were utilised in order to systematically and simultaneously vary all input parameters of the SRB algorithm efficiently (Unal et al., 1998). The purpose was to determine the most significant input parameters and

interactions that influence the global estimates for each region studied. Later a sensitivity analysis was conducted using Monte Carlo simulation to quantify the variability of the output data as the inputs changed within their range.

3 Sample space

Four global regions were selected for this study to include heavy foliage, pelagic, desert and mountainous in order to represent diverse atmospheric conditions and to correlate some latitude and longitude points with ground site locations as shown in Table 1. The ground site locations correspond to land surface measurements sites from the baseline surface radiation network (BSRN) (Ohmura et al., 1998) and the NOAA Pacific Marine Environment Laboratory (PMEL) group buoy sites (McPhaden et al., 1998). Future studies could use the results from this study for comparison with measurements obtained from these locations. There is no ground site located at the highest point of Mt. Everest. This point was selected to endorse and illustrate the advantage of satellite derived data as satellites can pass over the geographic coordinates for locations where measuring equipment positioning is not practical.

Table 1 Correlation of regions with ground sites

Region	*Location*	*LAT*	*LON*	*Ground site*
Amazon Rain Forest	Rolim de Moura, Brazil	–11.58	298.22	BSRN 73
Sahara Desert	Gobabeb, Namib Desert, Africa	23.56	15.04	BSRN 20
Indian Ocean	Indian Ocean	–7.97	67.00	BUOY RBJ
Mt. Everest	Highest Point	27.59	273.45	Not represented

4 SRB algorithm input parameters

The input data used for the SRB Version 4.0 algorithm are obtained from various sources, mostly from the International Satellite Cloud Climatology Project. Cosine solar zenith angle, cosine satellite zenith angle and azimuth angle are obtained from ISCCP for each satellite pixel. Cloud fraction, cloudy shortwave radiance, clear shortwave radiance and clear sky composite shortwave radiance are also averaged from ISCCP satellite observations. Precipitable water and column ozone amount are provided as part of the ancillary ISCCP datasets. Information about the surface such as the surface type and the snow/ice fraction are also provided in the ISCCP suite of data products. The SRB SW algorithm assumes a spectral variation of the surface Albedo which is fixed to this surface type but is modified when snow and/or ice is present. First guess aerosol optical depth and aerosol single scattering Albedo and aerosol asymmetry parameters are obtained from the MAC v1 dataset (Kinne et al., 2013). The products have been refined over the SRB versioning process. Ongoing issues include satellite calibration shifts due to orbital drift, small discontinuities during satellite transitions and uncertainties of various retrieved and averaged input quantities (GEWEX, 2015).

The 13 input parameters for the SRB algorithm and their short names are listed in Table 2. Table 2 also includes the full range of values for each parameter and their data sources.

Table 2 SRB algorithm input parameters and descriptions

Algorithm input variables			
Parameter assignment	*Description*	*Full range*	*Source*
colza	Cosine solar zenith angle	0.0 (sun on horizon)–1.0 (sun overhead)	ISCCP zenith angle
catza	Cosine satellite zenith angle	0.0 (satellite on horizon)–1.0 (satellite over cell)	ISCCP
Azi	Azimuth angle	0.0–180.0	ISCCP
cldfrc	Cloud fraction	0.0–1.0	ISCCP
cldrad	Cloudy shortwave radiance	0.0–1.11	ISCCP
clrrad	Clear shortwave radiance	0.0–1.11	ISCCP
cmprad	Clear sky composite shortwave radiance	0.0–1.11	ISCCP
pwater	Precipitable water	1.0–50.0	ISCCP
ozone	Column ozone	5.0–50.0	TOMS/TOVS/OMI (need ref)
pshcld	Phase of the cloud	1 = liquid 2 = ice	ISCCP
aertau	First guess aerosol optical depth	0.0–1.0	MAC v1
aerssa	Aerosol single scattering Albedo	0.9–1.0	MAC v1
aerasy	Aerosol asymmetry parameter	0.5–1.0	MAC v1

5 SRB output data for global estimates

The SRB shortwave algorithm has produced reliable data since its inception in 1994 and has been successfully used in science and industry. The output datasets for this study were produced using a modified version of the same code running the GEWEX SW algorithm, a modified version of an earlier physical model that derives surface solar irradiance from satellite observations (Pinker and Laszlo, 1991) and comprised of atmospheric properties calculations with radiative transfer. Where the Pinker/Laszlo code produces 3-hourly global data averaged from several points within a cell, the modified software for this study operates on a single latitude and longitude grid box (i.e., 1° × 1° area) and uses an average of one month of daily data. It produces 14 output parameters (global estimates) as shown in Table 3.

Table 3 SRB algorithm output data parameters

TOA (top-of-atmosphere) downward flux	W/m^2
TOA (top-of-atmosphere) upward flux	W/m^2
Surface downward flux (all-sky conditions)	W/m^2
Surface downward diffuse flux (all-sky conditions)	W/m^2
Surface upward flux (all-sky conditions)	W/m^2
Surface downward diffuse photosynthetically active radiation (PAR) (all-sky conditions)	W/m^2
Surface downward PAR (all-sky conditions)	W/m^2
TOA upward clear sky flux (no clouds)	W/m^2
Surface downward clear sky flux (no clouds)	W/m^2
Surface upward clear sky flux (no clouds)	W/m^2
Residual aerosol optical depth	Unitless
Residual cloud optical depth	Unitless
Surface downward pristine sky flux (no clouds, aerosols)	W/m^2
TOA upward pristine sky flux (no clouds, aerosols)	W/m^2

6 Experimental design

DOE enables efficient sampling of the design space and helps to determine the magnitude and relative importance of input parameters and their interactions (Box and Wilson, 1951). For this study, a determinant optimal experimental design (D-optimal) was constructed that enables the analysis of the 14 input parameters at three levels using only 128 experiments (Unal et al., 1999). A full factorial experimental design would have required about 1.6 million combinations (3^{13}). The modified version of the SRB algorithm enabled the global calculations of the algorithm to accept information for a single temporal and spatial point and for one month of averaged data. The points were from each of four atmospherically distinct regions to include the Amazon Rainforest, Sahara Desert, Indian Ocean and Mt. Everest. The same design was used for all regions.

In this study, there were 13 input parameters for the SRB algorithm using the satellite data, each having a defined range of values. The ranges were studied at three levels (low, medium and high), with the mean of the range representing the medium value. JMP® Statistical Software (JMP®, 2012) was used to construct a D-optimal design matrix of 128 experiments (rows) using coded values –1, 0, or 1 for the three levels (Unal et al., 1998). The coding enables mathematically independent assessment of input parameter effects. Table 4 shows a view of the coded design matrix used for each of the four regions.

The coded values (–1, 0 and 1) in each row of the D-optimal design were then converted to actual values within the ranges for each input parameter. Each of the 128 rows of the design represent a different combination of high, medium and low values of the SRB input parameters as shown in Table 5. There is a co-dependency among cosine of the solar zenith angle with clear sky radiance, cloudy sky radiance and clear sky composite shortwave radiance. This was an indication that real predictors would not extend to the limits of the range. The co-dependency was compensated for by

programmatically recognising an infeasible response during conditions of low sun and low satellite angles and adjusting the input as determined by the value of the cosine of the solar zenith angle (colza). These values are listed as VAR in Table 5.

Table 4 D-optimal design with coded values

Trial	*colza*	*catza*	*azi*	*cldfrc*	*cldrad*	*clrrad*	*cmprad*	*pwater*	*ozone*
1	1	1	–1	–1	–1	1	1	1	1
2	–1	0	–1	–1	1	–1	1	1	1
3	1	0	–1	1	–1	0	–1	1	0
4	–1	1	–1	1	0	1	–1	–1	1
5	–1	–1	0	1	1	0	1	–1	1
6	–1	–1	1	–1	–1	–1	1	–1	1
7	–1	1	–1	–1	1	1	1	–1	0
8	1	–1	–1	–1	–1	–1	–1	1	–1
9	1	1	–1	–1	1	–1	–1	–1	1
10	1	–1	1	–1	–1	1	–1	–1	–1
*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*
121	0	1	1	0	–1	–1	1	–1	–1
122	1	–1	–1	–1	0	0	0	1	1
123	1	1	1	1	0	1	1	1	1
124	–1	–1	–1	–1	–1	1	–1	0	–1
125	1	1	–1	1	–1	1	–1	–1	–1
126	1	–1	1	1	–1	–1	–1	1	1
127	0	0	1	–1	1	–1	–1	0	–1
128	–1	–1	1	0	1	1	0	–1	–1

Table 5 Actual input values corresponding to coded values of –1, 0 and +1

Colza	*Catza*	*Azi*	*Cldfrc*	*Cldrad*	*Clrrad*	*Cmprad*
0.20	0.32	0.01	0.00	0.07	0.05	0.05
0.58	0.66	89.51	0.50	VAR	VAR	VAR
0.96	0.99	179.00	1.00	0.60	0.36	0.33
Pwater	*Ozone*	*pshcld*	*aertau*	*aerssa*	*aerasy*	
13.48	24.21	1	0.07	0.89	0.58	
36.86	26.09	n/a	0.41	0.93	0.61	
60.25	27.96	2	0.76	0.97	0.64	

Table 6 shows the design matrix populated with the actual values for the Amazon Rainforest region.

The SRB algorithm also takes as input six additional static parameters. Spatial inputs such as latitude and longitude are set and kept constant for each region. One month of a

year was chosen for the temporal values for all regions to establish a fixed distance from the earth to the sun. The snow/ice unit is determined by region. The Mt. Everest region is assumed to be 1 and was set to 0 for the other three regions. The values for the static parameters are shown in Table 7.

Table 6 D-optimal design for the Amazon Rainforest grid box with actual values

1	0.96	0.99	0.01	0.00	0.20	0.36	0.33	60.25	27.96	2	0.76	0.97	0.61
2	0.20	0.66	0.01	0.00	0.20	0.05	0.12	60.25	27.96	1	0.07	0.97	0.64
3	0.96	0.66	0.01	1.00	0.20	0.30	0.24	60.25	26.09	1	0.07	0.97	0.64
4	0.20	0.99	0.01	1.00	0.14	0.13	0.05	13.48	27.96	1	0.07	0.97	0.64
5	0.20	0.32	89.51	1.00	0.20	0.09	0.12	13.48	27.96	2	0.07	0.89	0.58
6	0.20	0.32	179.00	0.00	0.07	0.05	0.12	13.48	27.96	1	0.07	0.97	0.58
7	0.20	0.99	0.01	0.00	0.20	0.13	0.12	13.48	26.09	2	0.07	0.97	0.58
8	0.96	0.32	0.01	0.00	0.20	0.24	0.24	60.25	24.21	2	0.76	0.89	0.64
9	0.96	0.99	0.01	0.00	0.60	0.24	0.24	13.48	27.96	1	0.07	0.97	0.58
10	0.96	0.32	179.00	0.00	0.20	0.36	0.24	13.48	24.21	1	0.76	0.97	0.64
*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*
118	0.20	0.66	0.01	1.00	0.20	0.09	0.12	60.25	24.21	2	0.07	0.89	0.61
119	0.20	0.66	179.00	1.00	0.14	0.13	0.12	36.86	24.21	1	0.07	0.89	0.58
120	0.20	0.32	89.51	0.00	0.20	0.05	0.12	13.48	24.21	2	0.76	0.97	0.61
121	0.58	0.99	179.00	0.50	0.16	0.14	0.26	13.48	24.21	1	0.41	0.89	0.61
122	0.96	0.32	0.01	0.00	0.40	0.30	0.29	60.25	27.96	2	0.07	0.97	0.58
123	0.96	0.99	179.00	1.00	0.40	0.36	0.33	60.25	27.96	2	0.07	0.89	0.64
124	0.20	0.32	10.00	0.00	0.07	0.13	0.10	36.86	24.21	2	0.41	0.97	0.58
125	0.96	0.99	0.01	1.00	0.20	0.36	0.24	13.48	24.21	1	0.76	0.93	0.58
126	0.96	0.32	179.00	1.00	0.20	0.24	0.24	60.25	27.96	2	0.07	0.97	0.61
127	0.58	0.66	179.00	0.00	0.60	0.14	0.14	36.86	24.21	1	0.76	0.89	0.64
128	0.20	0.32	159.00	0.50	0.20	0.13	0.10	13.48	24.21	2	0.76	0.93	0.58

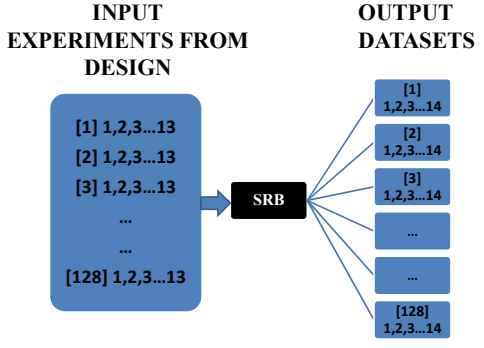
Table 7 Static input values for the SRB algorithm

Algorithm static input variables						
Region	*Month*	*Year*	*Latitude*	*Longitude*	*Satellite position ID*	*Snow/ice*
Amazon Rainforest	JUL	2007	–11.58	298.23	1	0
Indian Ocean	JUL	2007	–7.97	67.00	1	0
Mt Everest	JUL	2007	27.59	273.45	1	1
Sahara Desert	JUL	2007	23.56	15.04	1	0

7 SRB algorithm execution using the design matrix

To conduct the experiment for each region, the D-optimal experimental design was programmatically parsed to extract and pass one row of inputs to the SRB algorithm software. This process is represented in Figure 1. Executing and compiling the results was automated for efficiency, speed, scalability and reuse. A utility script created 128 NAMELIST files that contained a value for each of the 13 varied input parameters and the six static input parameters. After the SRB algorithm processed all 128 NAMELIST files, the utility script created a table of the output parameters to be used as dependent variables for the regression analysis. This was done for each of the four regions equating to 4 × 14 × 128 = 7,168 experiments.

Figure 1 A schematic showing the relationship between the input and output data parameters for the GEWEX SRB SW algorithm processing (see online version for colours)



8 Regression analysis and surrogate model building

Once all the trials prescribed by the design matrix were conducted using the SRB algorithm, a least squares regression analysis was performed to construct a second order mathematical model in the form of $Y = b_0 + \sum b_i x_i + \sum\sum b_{ij} x_{ij} + \sum b_{ii} x_i^2$ where b_0 is the Y intercept, Y is the dependent output variable value (Myers and Montgomery, 1995). Lower case bi values represent the coefficients from the regression analysis and x_i's represent one of the 13 input parameters. This mathematical model considers the effect of each input parameter on the global estimates, its significance, interaction effects and second order nonlinearity. Software was developed in order to rapidly build the 56 model equations for each of the 14 dependent variable values in each of the four regions. A separate regression was done for each dependent variable in all four regions.

9 Construct model equations

A second order math model, approximating the relationship between the dependent variables and each of the predictor variables in each of the four regions, was constructed for each dependent variable using the significant input parameters and interactions. The

purpose of constructing this model was to determine the most significant input parameters and interaction that influence the 14 global estimates in each region. These models were later used in a sensitivity analysis study using Monte Carlo simulation to quantify the variability of the output data as the inputs changed within their range.

As an example, math model coefficients indicating influence on the TOAUP estimate for each input parameter and interaction (factors) in the Amazon Rainforest region and the resulting mathematical model is shown on Table 8. Since all parameter and interaction values are scaled between –1 and +1, coefficient values listed in Table 8 are comparable in terms of influence. Larger the coefficient, larger is the influence.

Table 8 Significant parameters and interaction calculations for global estimate TOAUP

TOAUP	
Factor	*Coeff*
colza	154.1117419
cldfrc*cldrad	54.06427701
cldrad	65.56859697
cldfrc	60.57505477
colza*cldrad	43.0150125
colza*cldfrc	24.21200051
azi*azi	–65.6740312
colza*catza	–19.86112823
pwater*pshcld	–31.34222036
colza*colza	–54.42111832
azi*pwater	–15.3607226
catza	17.51297749
pwater	15.01052909
cmprad*aerasy	–9.782613058
catza*aertau	–10.07362949
aertau*aerssa	10.31365884
clrrad*clrrad	29.61860823
azi*pshcld	–17.40928998
clrrad	12.09283592
cmprad	11.68931142
colza*clrrad	8.94938664
cldfrc*aerasy	-8.72282627
colza*aertau	-8.010861627
cldfrc*aertau	-8.342771651
colza*azi	8.149323664

The model equation is built from the information in Table 8. The alpha character ‘A’ is a column in a spreadsheet followed by a number that represents the high, medium or low value for each predictor. These will be randomly varied during the sensitivity analysis. For this equation, Y is the dependent variable TOAUP.

$$Y = 316.505385 + 154.1117419 * A3 + 54.06427701 * A6 * A7 + 65.56859697$$
$$*A7 + 60.567505477 * A6 + 43.0150125 * A3 * A7 + 24.21200051 * A3 * A6$$
$$+ - 65.6740312 * A3 * A7 + 24.21200051 * A3 * A6 + -65.6740312 * A5 *$$
$$+ - 19.86112823 * A3 * A4 + -31.34222036 * A10 * A12 + -54.42111832 * A3$$
$$*A3 + -15.3607226 * A5 * A10 + 17.51297749 * A4 + 15.01052909 * A10$$
$$+ - 9.782613058 * A9 * A15 + -10.07362949 * A4 * A13 + 10.31365884 * A13$$
$$*A14 + 29.61860823 * A8 * A8 + -17.40928998 * A5 * A12 + 12.09283592 * A8$$
$$+11.68931142 * A9 + 8.94938664 * A3 * A8 + -8.72282627 * A6 * A15$$
$$+ - 8.010861627 * A3 * A13 + -8.342771651 * A6 * A13 + 8.149323664 * A3 * A5$$

10 Results

The analysis results indicate that the most significant parameter influencing the global estimates was colza. Colza interacted with other parameters differently in each region. These results are summarised in Table 9.

Table 9 Strongest input parameters and Interactions by region

Region	*Parameter*	*Interactions*
Amazon Rainforest	colza	colza * cldfrc
Sahara Desert	colza	colza * cldrad
Mt Everest	colza	colza * cldfrc
Indian Ocean	colza	cldfrc * cldrad

The data were also analysed to identify the global estimates that had the most predictors and predictor interactions with P-values less than 0.11. These are shown in Table 10.

Table 10 Dependent variables with highest number of influential predictors and predictor interactions by region

Region	*Dependent variable*
Amazon Rainforest	Srfdwnflx
Sahara Desert	Srfdwnprs
Mt. Everest	Srfdwnflx
Indian Ocean	Srfdwnpar

An analysis was performed to determine if there were significant input parameters and interactions that were common among the regions. The results are shown in Table 11. For each dependent variable in the first column of Table 11, input parameters and interactions that were found in all four regions are listed in the second column.

Table 11 Common significant input parameters for dependent variables among regions

Dependent variable	*Predictors and predictor interactions*
OAOD	Clrrad
OCOD	colza, catza * cldrad
SRFDWNCLRSKY	colza, clrrad
SRFDWNDIFF	colza, cldfrc * cldrad
SRFDWNDIFPAR	Colza
SRFDWNFLX	cldfrc * cldrad, cldfrc, colza, cldrad, cldfrc * pshcld, colza * cldfrc
SRFDWNPAR	cldrad, colza, cldfrc * cldrad, cldfrc * pshcld, colza * cldfrc
SRFDWNPRS	pshcld * aerasy, cldfrc * ozone, catza, pwater * pwater, colza, colza * catza, ozone * aertau, cmprad, aerasy, cldfrc * pshcld, colza * colza, pwater, cmprad * aerssa, ozone, colza * pwater
SRFUPCLRFLX	Colza
SRFUPFLX	colza * catza, cldrad, cldfrc
TOAUP	pwater * pshcld, azi * pwater, cldfrc * cldrad, cldfrc, cldrad, colza, colza * cldrad, catza * aertau, azi * pshcld, azi * azi, clrrad * clrrad, colza * cldfrc, clrrad
TOAUPCLRSKY	cmprad, colza, colza * catza, catza * azi
TOAUPPRS	cmprad, colza, colza * catza, colza * cmprad

Table 12 Results of sensitivity analysis

Variability				
Dependent variable	*Amazon*	*Sahara*	*Mt. Everest*	*Indian Ocean*
TOAUP	254	202.5	319	306
SRFDWNFLX	548	447	720	492
SRFDWNDIFF	221.5	277	116.2	317
SRFUPFLX	190.3	125.3	43.4	74.1
SRFDWNDIFPAR	55.8	155.1	60.5	158.1
SRFDWNPAR	258	211.5	272	242
TOAUPCLRSKY	174.3	114.8	43.9	87
SRFDWNCLRSKY	586	586	630	525
SRFUPCLRFLX	205.7	142.6	49.3	61.4
OAOD	0.488	0.977	0.057	0.613
OCOD	146.7	146.3	130.2	140.2
SRFDWNPRS	604	655	628	599
TOAUPPRS	199.5	139.7	52.1	160.3

11 Sensitivity analysis

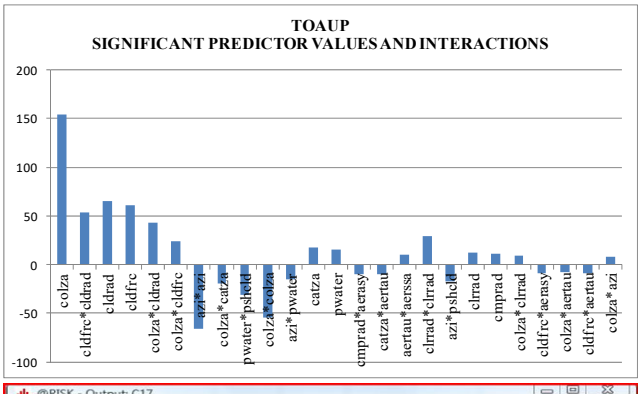
To determine the variability in the output data for each dependent variable varying within their range, a Monte Carlo simulation was conducted. This enabled more variations of predictor levels beyond the designed experiment to be evaluated. Triangular distributions

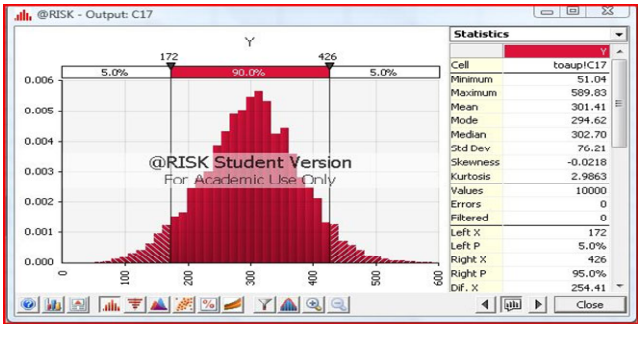
were constructed based on the lower (–1), medium (0) and high (+1) values for each input parameter range studied. The simulation varied all input parameters in the model equation about their normalised distribution ranges 10,000 times for each dependent variable and for each region. The resulting distribution curve of three standard deviations for each simulation along with statistical output quantified the variability. Table 12 indicates the variability for each dependent variable by region.

Figure 2 shows the graphical representation of the sensitivity analysis for the dependent variable TOAUP in the Amazon Rainforest region to include the coefficient table, a graph of the coefficient table and the variability histogram produced by @Risk® 4.5 software (Palisade Corporation, 2005).

Figure 2 Amazon region sensitivity analysis – TOAUP (see online version for colours)

TOAP	
Factor	***coeff***
colza	154.1117419
cldfrc*cldrad	54.06427701
cldrad	65.56859697
cldfrc	60.57505477
colza*cldrad	43.0150125
colza*cldfrc	24.21200051
azi*azi	-65.6740312
colza*catza	-19.86112823
pwater*pshcld	-31.34222036
colza*colza	-54.42111832
azi*pwater	-15.3607226
catza	17.51297749
pwater	15.01052909
cmprad*aerasy	-9.782613058
catza*aertau	-10.07362949
aertau*aerssa	10.31365884
clrrad*clrrad	29.61860823
azi*pshcld	-17.40928998
clrrad	12.09283592
cmprad	11.68931142
colza*clrrad	8.94938664
cldfrc*aerasy	-8.72282627
colza*aertau	-8.010861627
cldfrc*aertau	-8.342771651
colza*azi	8.149323664





12 Conclusions

This study was conducted to determine the most significant input parameters and their interactions in determining global estimates using the SRB algorithm based on satellite data. The analysis was conducted for four regions of the Earth. DOE methods were utilised for this purpose and a sensitivity analysis was conducted to quantify the variability in the input parameters for the SRB algorithm within the ranges studied. Second order response models were built using the resulting regression coefficients. A Monte Carlo simulation of each model added 10,000 additional experiments based on the

model developed and extended the probability distribution beyond the initial design trials to quantify variability in the modelled output data.

Results indicate that the cosine solar zenith angle was the strongest influence on the output data in all four regions. The interaction of cosine solar zenith angle and cloud fraction had the strongest influence on the output data in the Amazon, Sahara Desert and Mt. Everest Regions, while the interaction of Cloud Fraction and Cloudy Shortwave Radiance most significantly affected output data in the Indian Ocean region.

These results may provide significant insight into understanding the influence of input parameters for global estimates and can be useful in the development of the next generation SRB algorithm utilising the satellite data. In addition, a similar approach can be utilised in understanding factors about the atmospheres of other planets, such as Mars, that influence space navigation and human habitation.

Acknowledgements

The NASA GEWEX SRB project is supported through the Radiation Sciences program led by Dr. Hal Maring and Science Division at NASA Headquarters led by Dr. Jack Kaye.

References

Palisade Corporation (2005) *Advanced Risk Analysis for Spreadsheets*, @Risk 4.5, Ithaca, NY.

Barkstrom, B. and Smith, G.L. (1986) 'The earth radiation budget experiment: science and implementation', *Reviews of Geophysics*, Vol. 24, No. 2, pp.379–390, May.

Box, G.E.P. and Wilson, K.B. (1951) 'On the experimental attainment of optimum conditions', *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 13, No. 1, pp.1–45.

Earth Observing System (EOSWEB) (2017) *SRB Data and Information*, June [online] https://eosweb.larc.nasa.gov/project/srb/srb_table (accessed 20 June 2017).

Global Energy and Water Exchange Project (GEWEX) (2015) *Known Data Irregularities* [online] http://gewex-srb.larc.nasa.gov/common/php/SRB_known_issues.php (accessed 10 June 2015).

Hinkelman, L.M., Stackhouse Jr., P.W., Wielicki, B.A., Zhang, T. and Wilson, S.R. (2009) 'Surface insolation trends from satellite and ground measurements: comparisons and challenges', *Journal of Geophysical Research*, Vol. 114, pp.1–18, D00D20, August DOI: 10.1029/2008JD011004.

JMP® (2012) *Design User's Guide*, SAS Institute Inc, Cary, NC.

Kato, S., Loeb, N.G., Rutan, D.A., Rose, F.G., Sun-Mack, S., Miller, W.F. and Chen, Y. (2012) 'Uncertainty estimate of surface irradiances computed with MODIS, CALIPSO-, and CloudSat-derived cloud and aerosol properties', *Surv. Geophys*., January, DOI: 10.1007/s10712-012-9179-x.

Kinne, S., O'Donnel, D., Stier, P., Kloster, S., Zhang, K., Schmidt, H., Rast, S., Giorgetta, M., Eck, T.F. and Stevens, B. (2013) 'MAC-v1: a new global aerosol climatology for climate studies', *J. Adv. Model. Earth Syst.*, Vol. 5, pp.704–740, DOI: 10.1002/jame.20035.

McPhaden, M.J. et al. (1998) 'The tropical ocean-global atmosphere observing system: a decade of progress', *J. Geophys. Res.*, Vol. 103, No. 14, pp.169–240.

Myers, R.H. and Montgomery D.C. (1995) *Response Surface Methodology*, Wiley Interscience., Newyork, NY.

National Aeronautics and Space Administration (NASA) (2017a) *Prediction of Worldwide Energy Resource* [online] [online] https://power.larc.nasa.gov/ (accessed 4 August 2017).

National Aeronautics and Space Administration (NASA) (2017b) *SSE-GIS* [online] https://asdc-arcgis.larc.nasa.gov/sse/ (accessed 4 August 2017).

National Aeronautics and Space Administration (NASA) (2017c) *Surface Meteorology and Solar Energy*, a renewable energy resource website (release 6.0) [online] https://eosweb.larc.nasa.gov/sse/ (accessed 4 August 2017).

Ohmura, A. et al. (1998) 'Baseline surface radiation network (BSRN/WCRP): new precision radiometry for climate research', *Bulletin of the American Meteorological Society*, Vol. 79, No 10, pp.2115–2136.

Pinker, R.T. and Laszlo, I. (1991) 'Effects of spatial sampling of satellite data on derived surface solar irradiance', *American Meteorological Society*, Vol. 8, No. 97, February, DOI: https://doi.org/10.1175/1520-0426(1991)008<0096:EOSSOS>2.0.CO;2.

Raschke, E., Kinne, S., Rossow, W.B., Stackhouse Jr., P.W. and Wild, M. et al. (2016) 'Comparison of radiative energy flow in observationl datasets and climate modeling', *Journal of Applied Meteorology and Climatology*, Vol. 55, No. 1, pp.93–117, January, DOI: http://dx.doi.org/10.1175/JAMC-D-14-0281.1

Rossow, W.B. and Zhang, Y-C. (1995) 'Calculation of surface and top-of-atmosphere radiative fluxes from physical quantities based on ISCCP datasets, part II: validation and first results', *Journal of Geophysical Research: Atmospheres*, Vol. 100, No. D1, pp.1167–1197.

Stackhouse Jr., P.W., Gupta, S.K., Cox, S.J., Zhang, T., Colleen Mikovitz, J. and Hinkelman, L.M. (2011) 'The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset', *GEWEX News*, Vol. 21, No. 1, pp.10–12, February.

Stephens, G.L, Li, J., Wild, M., Clayson, C.A., Loeb, N., Kato, S., L'Ecuyer, T., Stackhouse, P.W. and Andrews, T.A. (2012) 'An update on earth's energy balance in light of the latest global observations', *Nature Geoscience*, Vol. 5, pp.691–696, DOI: 10.1038/ngeo1580 [online] http://www.nature.com/ngeo/journal/v5/n10/abs/ngeo1580.html (accessed 12 July 2015).

Unal, R, Lepsch, R.A. and McMillin, M.L. (1998) 'Response surface model building and multidisciplinary optimization using D-optimal designs', *7th Annual AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Paper No: AIAA-98-4759, September.

Unal, R., Lepsch, R.A., Lockwood, Mary K. and McMillin, M.L. (1999) 'Model building for rapid multidisciplinary integration and optimization using experimental designs', *Proceedings of the 20th ASEM National Conference*, Virginia Beach VA, October, pp.253–259.

Whitlock, C.H., Charlock, T.P., Staylor, W.F., Pinker, R.T., Laszlo, I., Ohmura, A., Gilgen, H. Konzelman, T., DiPasquale, R.C., Moats, C.D., LeCroy, S.R. and Ritchey, N.A. (1994) *Bulletin of the American Meteorological Society*, 75th Annual Meeting Issue, June, Vol. 76, No. 6, pp.905–922.

Zhang, T., Stackhouse Jr., P.W., Gupta, S.K., Cox, S.J., Mikovitz, J.C. and Hinkelman, L.M. (2013) 'The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: a systematic quality control, production and application approach', *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 122, pp.127–140, June [online] http://www.sciencedirect.com/science/article/pii/S0022407312004335 (accessed 31 July).

Zhang, T., Stackhouse Jr., P.W., Gupta, S.K., Cox, S.J., Mikovitz, J.C. and Hinkelman, L.M. (2015) 'The validation of the GEWEX SRB surface longwave flux data products using BSRN measurements', *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 150, pp.134–147, January [online] http://www.sciencedirect.com/science/article/pii/S0022407312004335 (accessed 31 July).

Zhang, Y., Rossow, W.B. and Stackhouse Jr., P.W. (2007) 'Comparison of different global information sources used in surface radiative flux calculation: radiative properties of the surface', *Journal of Geophysical Research*, Vol. 112, pp.1–20, D01102, January, DOI: 10.1029/2005JD007008.