Supplement of Geosci. Model Dev., 12, 3975–3990, 2019
https://doi.org/10.5194/gmd-12-3975-2019-supplement

Geoscientific Model Development
Open Access

Supplement of

Developing a monthly radiative kernel for surface albedo change from satellite climatologies of Earth's shortwave radiation budget: CACK v1.0

Ryan M. Bright and Thomas L. O'Halloran

Correspondence to: Ryan M. Bright (ryan.bright@nibio.no)

Machine learning results summary

A subset of the machine learning model solutions for ECHAM6 and CAM5 and associated summary statistics are presented in Table S1. Equivalent solutions of complexities 1, 6, 7, and 10 were found independently by the two GCMs. The model with complexity 10 (red boldface) was the highest performing model common to both GCMs and was chosen to represent $K_{\alpha_s}^{BO18}$ and hence subjected to further performance evaluation in the main article.

Table S1. Subset of machine learning model solutions for ECHAM6 and CAM5 and associated statistics for the "selected" solutions shown in Figure S1. Means of CAM5 and ECHAM6 kernels are 140.2 and 133.4 W m^{-2}, respectively, which are used to compute the monthly relative RMSEs ("rRMSE"; in %). "Comp." = model complexity.

Solution	**R²**	**Max Error**	**rRMSE (%)**	**MSE**	**MAE**	**Comp.**
ECHAM6						
$K=\left[SW_{\uparrow}^{SFC}+SW_{\uparrow}^{TOA}+SW_{\downarrow}^{SFC}\sqrt{SW_{\downarrow}^{SFC}}-SW_{\downarrow}^{TOA}\right]/\sqrt{SW_{\downarrow}^{TOA}}+\log\left(SW_{\downarrow}^{TOA}+SW_{\downarrow}^{TOA}SW_{\downarrow}^{SFC}\right)$	1.00	31.96	4.1	27.02	3.92	30
$K=\left[SW_{\uparrow}^{SFC}+SW_{\downarrow}^{SFC}\sqrt{SW_{\downarrow}^{SFC}}\right]/\sqrt{SW_{\downarrow}^{TOA}}$	0.99	32.59	5.6	40.93	4.91	16
$K=SW_{\downarrow}^{SFC}\sqrt{SW_{\downarrow}^{SFC}/SW_{\downarrow}^{TOA}}$	**0.99**	**31.59**	**7.1**	**66.48**	**6.03**	**10**
$K=SW_{\downarrow}^{SFC}-\sqrt{SW_{\downarrow}^{TOA}}$	0.86	81.55	22.0	947.8	25.10	7
$K=SW_{\downarrow}^{SFC\,2}/SW_{\downarrow}^{TOA}$	0.81	82.95	31.4	1,314	33.03	6
$K=SW_{\downarrow}^{SFC}$	0.67	103.8	41.0	2,245	40.98	1
CAM5						
$K=\sqrt{SW_{\uparrow}^{SFC}\sqrt{SW_{\uparrow}^{SFC}}+SW_{\downarrow}^{SFC\,3}/SW_{\downarrow}^{TOA}+SW_{\uparrow}^{SFC}\sqrt{SW_{\downarrow}^{TOA}+SW_{\uparrow}^{SFC}}}$	1.00	43.57	5.0	35.66	4.31	30
$K=SW_{\downarrow}^{SFC}+\left(SW_{\uparrow}^{SFC}-SW_{\downarrow}^{TOA}\right)/\log\left(SW_{\downarrow}^{TOA}\right)+\log\left(SW_{\downarrow}^{TOA}\right)$	0.99	53.39	5.7	45.82	5.31	18
$K=SW_{\downarrow}^{SFC}\sqrt{SW_{\downarrow}^{SFC}/SW_{\downarrow}^{TOA}}$	**0.99**	**36.62**	**7.7**	**83.37**	**6.71**	**10**
$K=SW_{\downarrow}^{SFC}-\sqrt{SW_{\downarrow}^{TOA}}$	0.88	82.26	25.0	874.9	23.87	7
$K=SW_{\downarrow}^{SFC\,2}/SW_{\downarrow}^{TOA}$	0.80	83.99	32.4	1,474	35.04	6
$K=SW_{\downarrow}^{SFC}$	0.71	103.9	38.7	2,098	39.03	1
ECHAM6 & CAM5 mean						
$K=K_{\alpha_s}^{BO18}=SW_{\downarrow}^{SFC}\sqrt{SW_{\downarrow}^{SFC}/SW_{\downarrow}^{TOA}}$	***0.99***	***34.11***	***7.4***	***74.93***	***6.37***	***10***

The rRMSE for $K_{\alpha_s}^{BO18}$ of Table S1 is the mean rRMSE for the ECHAM6 and CAM5 solutions.

Figure S1 illustrates the Pareto front used to assist $K_{\alpha_s}^{BO18}$ model selection. Model solutions are plotted as small dots showing model MSE as a function of model complexity. A subset of models of interest, generally found at 'elbows' in the Pareto front, are indicated by larger dots. At these elbows, slight increases in model complexity lead to large reductions in model error. $K_{\alpha_s}^{BO18}$ has a model complexity of 10.

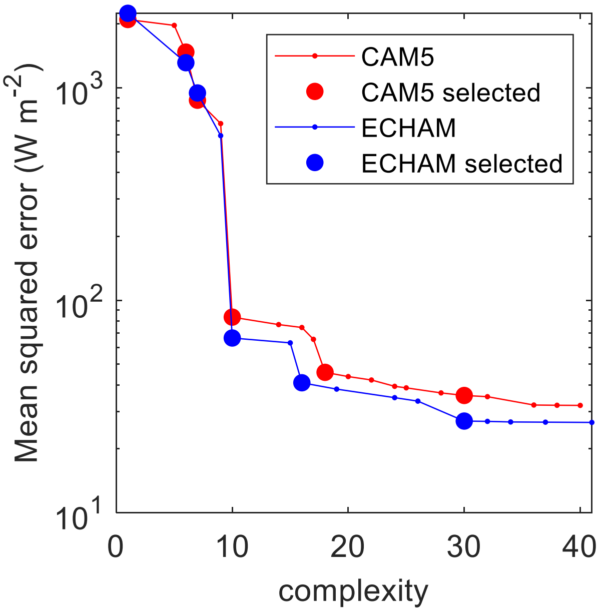

Figure S1. Pareto front used to assist model selection from machine learning output.

Additional uncertainty detail

The rRMSE for $K_{\alpha_s}^{BO18}$ (Table S1) is used to estimate CACK's monthly model error for a given grid cell as follows:

$$\sigma_{me,m,p}(CACK) = \frac{RMSE}{\left|\overline{K_{\alpha_s}^{GCM}}\right|} CACK_{m,p} \tag{S1}$$

where the numerator represents the mean RMSE of the ECHAM6 and CAM5 solutions, $\overline{K_{\alpha_s}^{GCM}}$ is the mean of the monthly ECHAM6 and CAM5 kernels in the training datasets, and $CACK_{m,p}$ is the CERES albedo change kernel (CACK) based on the $K_{\alpha_s}^{BO18}$ parameterization for month *m* and grid cell *p*.

Uncertainty related to the local "physical variability" of a CERES input variable x for any given month *m* and grid cell *p* is taken as the standard deviation within the 2001-2016 period:

$$\sigma_{PV,m,p}(x) = \sqrt{\frac{\sum \left| x_{m,p} - \bar{x}_{m,p} \right|}{Y}} \tag{S2}$$

where $x_{m,p}$ is the monthly and grid cell value of x, Y is the total number of years in the period (i.e., 16), and $\bar{x}_{m,p}$ is the monthly and grid cell mean of x during this period.

Uncertainty related to CERES EBAF input variable x for any given month *m* and grid cell *p* is estimated using its relative uncertainty as:

$$\sigma_{DU,m,p}(x) = \frac{\sigma(x)}{\left| \bar{x} \right|} x_{m,p} \tag{S3}$$

where x_m is the monthly and grid cell value of x, $\sigma(x)$ is the absolute uncertainty of x (Table S2), and $\bar{x}$ is the mean x of the sample domain (Table S2).

Table S2. Uncertainty of the CERES EBAF v4 input variables required by CACK.

CERES Variable	**Domain**	x (W m^{-2})	$\sigma(x)$ (W m^{-2})	$\sigma(x)/\left\|\bar{x}\right\|$	**Reference**
$SW_{\downarrow}^{SFC}$	"Ocean + land"	187	13	0.07	(Kato et al., 2018)
$SW_{\downarrow}^{TOA}$	N/A	N/A	N/A	0.01	(Loeb et al., 2017)

Covariance of CERES input variables x_1 and x_2 (i.e., $SW_{\downarrow}^{SFC}$ and $SW_{\downarrow}^{TOA}$) in any given grid cell *p* is estimated as:

$$\sigma_p(x_1, x_2) = \frac{1}{M-1} \sum_{m=1}^{M} (x_{1,m,p} - \bar{x}_{1,p})(x_{2,m,p} - \bar{x}_{2,p}) \tag{S4}$$

where M is the total number of months, $x_{1,m,p}$ and $x_{2,m,p}$ are the values for variables x_1 and x_2 in grid cell *p* and month *m*, and $\bar{x}_{1,p}$ and $\bar{x}_{2,p}$ are the means of x_1 and x_2 in grid cell *p* for the 2001-2016 time period.

Grid cell σ and grid cell and monthly σ_{PV} and σ_{DU} are then used to estimate the total propagated uncertainty of CERES input variables x_1 and x_2 (i.e., $SW_{\downarrow}^{SFC}$ and $SW_{\downarrow}^{TOA}$) as the second right-hand term of Eq. (21) of the main article. This is then summed together with the σ_{me} estimated as equation S1. In Eq. (21), the partial derivative of CACK with respect to $SW_{\downarrow}^{SFC}$ is given as:

$$\frac{\partial CACK}{\partial SW_{\downarrow}^{SFC}} = \frac{3\sqrt{SW_{\downarrow}^{SFC} / SW_{\downarrow}^{TOA}}}{2} \tag{S5}$$

The partial derivative of CACK with respect to $SW_{\downarrow}^{TOA}$ is given as:

$$\frac{\partial CACK}{\partial SW_{\downarrow}^{TOA}} = -0.5\left(\frac{SW_{\downarrow}^{SFC}}{SW_{\downarrow}^{TOA}}\right)^{\frac{3}{2}} \quad \text{(S6)}$$

CACK v1.0 dataset summary

Table S3 summarizes the variables comprising the CACK v1.0 dataset.

Table S3. Summary of variables included in the CACK v1.0 dataset.

Variable name	Description	Temporal resolution	Temporal signature (extent)
"CACK CM"	Eq. (17) with CERES inputs as 2001-2016 means	Monthly	2001-2016 mean
"Sigma_me CM"	First right-hand term of Eq. (21) estimated as Eq. (S1) with CERES inputs as 2001-2016 means	Monthly	2001-2016 mean
"Sigma_du_pv CM"	Second right-hand term in Eq. (21) where $\sigma(x_1, x_2)$, $\sigma_{PV}(x_n)$, and $\sigma_{DU}(x_n)$ are estimated with Eq.'s S2-S4, and x_1 an d x_2 are 2001-2016 means of $SW_{\downarrow}^{SFC}$ and $SW_{\downarrow}^{TOA}$	Monthly	2001-2016 mean
"Sigma_total CM"	Eq. (21), or the sum of "Sigma_me CM" and "Sigman_du_pv CM"	Monthly	2001-2016 mean
"CACK"	Eq. (17) estimated for all years in the 2001-2016 period	Monthly	Annual (2001-2016)
"Sigma_me"	First right-hand term of Eq. (21) estimated as Eq. (S1) for all years in the 2001-2016 period	Monthly	Annual (2001-2016)
"Sigma_du"	Second right-hand term of Eq. (21) estimated for all years in the 2001-2016 period but excluding $\sigma_{PV}(x_n)$; $\sigma(x_1, x_2)$ and $\sigma_{DU}(x_n)$ are estimated with Eq.'s S3-S4; x_1 an d x_2 are $SW_{\downarrow}^{SFC}$ and $SW_{\downarrow}^{TOA}$ provided by CERES EBAF v4.	Monthly	Annual (2001-2016)
"Sigma_total"	Sum of "Sigma_me" and "Sigma_du"	Monthly	Annual (2001-2016)

References

Kato, S., Rose, F. G., Rutan, D. A., Thorsen, T. J., Loeb, N. G., Doelling, D. R., Huang, X., Smith, W. L., Su, W., and Ham, S.-H.: Surface Irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Data Product, Journal of Climate, 31, 4501-4527, 10.1175/JCLI-D-17-0523.1, 2018.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and Kato, S.: Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product, Journal of Climate, 31, 895-918, 10.1175/JCLI-D-17-0208.1, 2017.