Supporting information

Predicting acidification recovery at the Hubbard Brook Experimental Forest, New Hampshire, USA: Evaluation of four models

Koji Tominaga†, Julian Aherne†, Shaun A. Watmough†, Mattias Alveteg‡, Bernard J. Cosby§, Charles T. Driscoll#, Maximilian Posch⊥, and Afshin Pourmokhtarian#

† Environmental and Life Sciences, Trent University, 1600 West Bank Drive, Peterborough, Ontario, K9J 7B8

‡ Department of Chemical Engineering, Lund University, P.O. Box 124, Lund 221 00, Sweden

§ Environmental Sciences Department, University of Virginia, 291 McCormick Road, Charlottesville, Virginia, 22904-4123, USA

Department of Civil and Environmental Engineering, Syracuse University, Syracuse, New York, 13244-1190, USA

⊥Coordination Centre for Effects (CCE), National Institute for Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands

Model configuration and calibration:

Table S1. Description of MAGIC, PnET-BGC, SAFE and VSD and their configuration in the current study.

Table S2. Model calibration success rate against behavioral criteria.

Model configurations. Key features are described in detail in Table S1, along with the configurations used in the current study.

Table S1. Model description of MAGIC, PnET-BGC, SAFE and VSD; configurations for the current study are given in parentheses (if ambiguous).

Features	MAGIC	PnET-BGC	SAFE	VSD
Number of soil layers	Up to 3 (1)	2 (organic and mineral)	Up to 20 (1)	1
Temporal resolution	Annual, monthly (annual)	Monthly	Annual	Annual
Base cation weathering	Calibrated, or external-input (external input)	External input	Sub-model or external input (external input)	External input
Forest nutrient uptake	External input	Sub-model	External input	External input
Runoff to precipitation ratio	External input (proportional to precipitation)	External input (proportional to precipitation)	Fixed ratio	External input (proportional to precipitation)
Sulfate adsorption	Langmuir isotherm	Yes	Yes (not used)	Not included
N immobilization	Fractional, fixed, or external input (fractional)	Sub-model	Available in previous versions (not used in this study)	Fractional, fixed, or external input (fractional)
Nitrification	Fractional (100 %)	Sub-model	Always, or own sub-model (fractional 100 %)	100 %
Denitrification	Fractional (none)	Sub-model	None or own sub-model (none)	Fractional, fixed (none)
Soil N build-up controlled by C:N ratio	Yes (not used)	Sub-model	No	Yes (not used)
CO_2 degassing in surface water	Yes	Yes	Yes	Yes
$Al(OH)_3$ precipitation in stream	Yes	No	Yes	Yes
Lumped base cations (Ca^{2+}, Mg^{2+}, and K^+)	No	No	Yes	Yes
DOC dissociation model	Triprotic	Triprotic	Oliver	Oliver, or simple monoprotic (Oliver)
DOC-Al complexation	Yes	Yes	No	No
Cation exchange	Gaines-Thomas	Gaines-Thomas	Gapon	Gaines-Thomas or Gapon (Gapon)
Cation exchange equations	4	6	1	2

Ions in soil solution charge balance	28	29	16	12
First appearance in the peer-reviewed literature	1985	1991	1993	2009

Model calibration. The calibration success rate, i.e., percentage of simulations (n = 10000) that met the calibration criteria (Table 2), varied by model and simulated variable (Table S2). The low success rate for stream Na^+ concentration simulated by MAGIC (11.8 %) compared with the other models (71.3–76.3 %) suggests that the initial parameter distributions, in particular, Na^+ weathering and initial exchangeable Na^+ fraction, were poorly defined for MAGIC (Table 1); stream [Na^+] is primarily characterized by Na^+ deposition, exchange on soil surface complex (VSD and SAFE do not model Na^+ exchange, Table S1) and weathering, but is also influenced by a variety of factors such as acidic deposition flux and soil solution parameters (e.g., dissolved organic carbon concentration, pCO_2, and gibbsite dissolution constant). These factors are modeled differently by each model.

Table S2. Success rates (%) for calibration criteria by simulated variables and model (n = 10000). Model simulations that were successful for all variables were considered behavioral (see Table 2).

Variable	MAGIC	PnET-BGC	SAFE	VSD
Soil base saturation (%)*	17.4	43.4	50.3	64.9
Surface water pH	15.1	30.7	36.1	17.4
Surface water [Bc (Ca^{2+} + Mg^{2+} + K^+)] (μeq L^{-1})	11.4	49.5	21.2	5.1
Surface water [Na^+] (μeq L^{-1})	11.8	71.3	76.3	76.3
Surface water [SO_4^{2-}] (μeq L^{-1})	25.9	35.4	49.1	49.6
Surface water [ANC] (μeq L–1)	12.5	27.2	29.5	18.2

All variables	0.3	0.4	1.7	0.5

* Base saturation is the percentage of the cation exchange capacity occupied by base cations (Ca^{2+} + Mg^{2+} + K^{+})

Another notable instance is the low success rate of VSD for stream [Bc = Ca^{2+} + Mg^{2+} + K^{+}] (5.1 %), again indicating poor initial parameter distribution compatibility for VSD with respect to the simulation of stream [Bc]. Similar to [Na^{+}], stream [Bc] is characterized by a series of factors, and the formulations of these processes differ substantially between models. Moreover, despite only a slight difference in the cation exchange formulation between SAFE [3] and VSD [4] in this study (Table 1), a substantial difference in the stream [Bc] success rate for SAFE (21.2 %) and VSD (5.1 %) was observed.

In the current study, the success rates for individual chemical variables were relatively high for PnET-BGC and SAFE, but their success rate for all variables (i.e., satisfying all criteria, 0.4 % and 1.7 %, respectively) was as low as for the other models (MAGIC, 0.3 % and VSD, 0.5 %). The behavioral parameter sets were limited by their initial distributions, and therefore the success rates are not indicative of each model's true calibration efficiency.