Water Resources Research

Supporting Information for

Large-scale groundwater monitoring in Brazil assisted with satellite-based artificial intelligence techniques

Clyvihk Renna Camacho[1,2], Augusto Getirana*[3,4], Otto Corrêa Rotunno Filho[2], Maria Antonieta A. Mourão[1]

1 Geological Survey of Brazil, Belo Horizonte, Brazil
2 Civil Engineering Department, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
3 Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD
4 Science Applications International Corporation, Greenbelt, MD

*Corresponding author (augusto.getirana@nasa.gov)

Contents of this file

Supporting Table S1, S2 and S3

Supporting Table S1. Effective porosity in Brazilian aquifers.

Aquifer	Effective porosity (n_e)	Reference
Açu	0.10	(Diniz et al., 2012)
Alter do Chão	0.18	(Aguiar & Mourão, 2012)
Areado	0.05	Estimated
Barreiras (Pirabas e Grajau)	0.10	(Silva et al., 2008)
Caiuá	0.17	(Franzini, 2012)
Beberibe	0.10	(Silva et al., 2008)
Boa Vista	0.18	Estimated
Cabeças	0.03	(Correia Filho et al., 2010)
Cenozoic Covers Aquifer	0.10	Estimated
Furnas	0.13	Estimated
Guarani	0.18	(Takahashi, 2012)
Içá	0.10	(Galvão et al., 2012)
Itapecuru	0.13	(Santos, 2005)
Coastal	0.14	Estimated
Missão/Velha-Mauriti	0.10	(Souza & Castro, 2013)
Parecis	0.15	(Silva, 2013)
Poti Piauí	0.15	(Correia Filho et al., 2010)
Prosperança	0.15	Estimated
Ronuro	0.12	(Peixoto et al., 2012)
Salto das Nuvens	0.13	(Peixoto et al., 2012)
Serra do Tucano	0.17	Estimated
Serra Grande	0.03	(Aguiar, 2017)
Tacaratu	0.03	(Diniz et al., 2012)
Trombetas	0.15	Estimated
Tucunaré	0.10	Estimated
Urucuia	0.13	(Gaspar & Campos, 2007)

Supporting Table S2. Sensitivity test of the models in Experiment E2. RMSE [cm], MAE [cm], NSE [-], KGE [-]. Extreme Gradient Boosting (XGB). Light Gradient Boosting Model (LGBM). CatBoost Model (CtB). Random Forest (RF). Ordinary Least Squares Model (OLS). Linear Regression (LR). Bayesian Ridge Model (BR). Stochastic Gradient Descent Model (SGDRegressor). Support-vector Machine (SVM). Multi-Layer Perceptron (MLP). Long Short-Term Memory (LSTM). The acronyms in parentheses indicate that the results of the models were used as input to the external model. In red the best values for the metric.

Calibration							
Model	**NSE**	**RMSE**	**MAE**	**KGE**	***r***	**α**	**β**
BR	0.10	4.10	2.71	-0.02	0.32	0.33	1.36
BR[LGBM, CtB]	0.45	3.23	2.12	-0.45	0.67	0.69	-0.38
BR[XGB, CtB]	0.57	2.81	1.69	0.29	0.76	0.75	1.62
BR[XGB, LGBM, CtB]	0.61	2.71	1.65	0.38	0.78	0.78	0.46
BR[XGB, LGBM]	0.60	2.77	1.70	0.64	0.77	0.76	0.85
CtB	0.24	3.77	2.39	0.05	0.49	0.47	1.60
CtB[LGBM, CtB]	0.46	3.24	2.10	0.37	0.68	0.65	1.41
CtB[XGB, CtB]	0.56	2.81	1.71	-7.45	0.75	0.72	9.44
CtB[XGB, LGBM, CtB]	0.58	2.79	1.70	0.63	0.76	0.73	1.04
CtB[XGB, LGBM]	0.59	2.81	1.68	-5.66	0.77	0.76	-5.65
LGBM	0.24	3.79	2.39	0.09	0.49	0.46	1.53
LGBM[LGBM, CtB]	0.42	3.19	2.03	0.46	0.65	0.59	0.96
LGBM[XGB, CtB]	0.60	2.80	1.72	-8.41	0.78	0.70	8.41
LGBM[XGB, LGBM, CtB]	0.58	2.83	1.74	0.43	0.76	0.72	1.43
LGBM[XGB, LGBM]	0.56	2.85	1.75	0.56	0.75	0.73	1.24
LR	0.09	4.12	2.72	-0.03	0.31	0.34	1.39
LR[LGBM, CtB]	0.44	3.25	2.12	0.32	0.66	0.69	0.50
LR[XGB, CtB]	0.57	2.73	1.63	0.47	0.76	0.77	0.59
LR[XGB, LGBM, CtB]	0.61	2.68	1.63	0.63	0.78	0.79	0.83
LR[XGB, LGBM]	0.44	3.28	2.13	-0.88	0.67	0.68	2.82
LSTM	0.15	4.06	2.59	-0.23	0.41	0.52	0.03
LSTMBidirectional	0.12	4.14	2.66	-0.31	0.37	0.49	-0.03
MLP	0.10	4.19	2.69	0.06	0.32	0.36	1.13
MLP[LGBM, CtB]	0.49	3.18	2.03	0.04	0.70	0.73	0.13
MLP[XGB, CtB]	0.51	3.13	1.95	-3.80	0.72	0.72	5.78
MLP[XGB, LGBM, CtB]	0.61	2.70	1.64	0.25	0.78	0.76	0.32
MLP[XGB, LGBM]	0.49	3.21	2.04	-0.06	0.70	0.77	0.01
OLS	0.09	4.12	2.72	-0.38	0.31	0.34	1.99
RF	0.19	3.89	2.43	0.26	0.46	0.59	1.28
RF[LGBM, CtB]	0.38	3.43	2.15	0.05	0.62	0.73	0.17
RF[XGB, CtB]	0.44	3.30	1.95	0.49	0.67	0.70	1.33

RF[XGB, LGBM, CtB]	0.41	3.47	2.04	-0.98	0.65	0.78	-0.94
RF[XGB, LGBM]	0.44	3.19	1.93	0.46	0.68	0.83	0.60
SGDRegressor	0.05	4.23	2.78	-0.16	0.25	0.38	1.62
SVM	0.11	4.08	2.55	-4.33	0.36	0.25	6.24
XGB	0.17	3.94	2.45	0.32	0.45	0.63	1.16
XGB[LGBM, CtB]	0.38	3.44	2.19	0.05	0.61	0.62	0.22
XGB[XGB, CtB]	0.44	3.33	1.97	-0.06	0.67	0.61	0.07
XGB[XGB, LGBM, CtB]	0.45	3.21	1.99	-0.75	0.69	0.53	-0.66
XGB[XGB, LGBM]	0.48	3.10	1.89	0.36	0.69	0.74	0.50
Validation							
Model	**NSE**	**RMSE**	**MAE**	**KGE**	***r***	**α**	**β**
BR	0.11	4.73	3.14	-0.08	0.33	0.28	0.54
BR[LGBM, CtB]	0.24	4.36	2.83	0.31	0.50	0.57	0.82
BR[XGB, CtB]	0.17	4.54	2.90	0.23	0.45	0.62	0.61
BR[XGB, LGBM, CtB]	0.17	4.55	2.90	0.15	0.45	0.62	0.48
BR[XGB, LGBM]	0.18	4.53	2.89	0.18	0.45	0.62	0.51
CtB	0.22	4.40	2.80	0.09	0.48	0.40	0.56
CtB[LGBM, CtB]	0.24	4.35	2.82	0.31	0.50	0.56	0.90
CtB[XGB, CtB]	0.20	4.47	2.88	0.21	0.47	0.61	0.57
CtB[XGB, LGBM, CtB]	0.20	4.47	2.89	0.19	0.47	0.62	0.52
CtB[XGB, LGBM]	0.19	4.49	2.89	0.20	0.47	0.62	0.54
LGBM	0.24	4.36	2.78	0.17	0.50	0.40	0.73
LGBM[LGBM, CtB]	0.23	4.37	2.81	0.23	0.48	0.50	0.74
LGBM[XGB, CtB]	0.19	4.50	2.90	0.21	0.46	0.60	0.60
LGBM[XGB, LGBM, CtB]	0.19	4.48	2.88	0.12	0.46	0.59	0.43
LGBM[XGB, LGBM]	0.18	4.52	2.92	0.07	0.45	0.61	0.36
LR	0.10	4.73	3.14	-0.02	0.33	0.29	0.70
LR[LGBM, CtB]	0.17	4.54	2.90	0.17	0.45	0.62	0.17
LR[XGB, CtB]	0.17	4.54	2.90	0.18	0.45	0.63	0.52
LR[XGB, LGBM, CtB]	0.24	4.35	2.50	0.31	0.50	0.57	0.84
LR[XGB, LGBM]	0.24	4.35	2.83	0.32	0.50	0.57	0.83
LSTM	0.06	4.85	3.16	-1.58	0.26	0.22	3.34
LSTMBidirectional	0.05	4.88	3.12	-0.21	0.24	0.13	0.62
MLP	0.06	4.85	3.23	-1.85	0.27	0.26	-1.65
MLP[LGBM, CtB]	0.14	4.91	2.82	0.07	0.49	0.53	0.73
MLP[XGB, CtB]	0.15	4.80	2.87	0.01	0.47	0.52	1.79
MLP[XGB, LGBM, CtB]	0.18	4.71	2.95	0.07	0.35	0.69	0.40
MLP[XGB, LGBM]	0.12	4.47	2.80	0.09	0.50	0.66	1.37
OLS	0.10	4.73	3.14	0.01	0.32	0.29	0.89
RF	0.21	4.45	2.83	0.08	0.46	0.50	0.45
RF[LGBM, CtB]	0.21	4.43	2.88	0.25	0.48	0.62	0.60

RF[XGB, CtB]	0.17	4.56	2.95	0.15	0.46	0.67	0.42
RF[XGB, LGBM, CtB]	0.16	4.58	2.95	0.15	0.46	0.67	0.43
RF[XGB, LGBM]	0.14	4.64	3.00	0.04	0.44	0.68	0.28
SGDRegressor	0.11	4.73	3.14	-0.08	0.33	0.28	0.54
SVM	0.09	4.76	3.01	-0.48	0.33	0.22	2.06
XGB	0.21	4.44	2.82	0.14	0.47	0.54	0.51
XGB[LGBM, CtB]	0.21	4.43	2.84	-0.02	0.47	0.53	0.27
XGB[XGB, LGBM]	0.18	5.10	2.95	-0.39	0.43	0.50	-0.17
XGB[XGB, CtB]	0.21	4.45	2.84	-0.31	0.46	0.52	-0.10
XGB[XGB, LGBM, CtB]	0.20	4.47	2.85	-0.65	0.45	0.45	-0.46

Supporting Table S3 - Summary of results for experiments E3 and E4 over Brazilian aquifers.

Aquifer	Experiment	Wells	Correlation RIMAS x GRACE	RMSE [cm]		MAE [cm]		NSE	KGE ΔGWS_{SIM}			
				$\Delta GWS_{S\,IM}$	$\Delta GWS_{CL\,SM}$	$\Delta GWS_{S\,IM}$	$\Delta GWS_{CL\,SM}$	$\Delta GWS_{S\,IM}$	KGE	*r*	α	β
Guarani	E3	Calib.	-0.30	1.65	3.63	1.43	2.8	0.72	0.54	0.86	0.82	0.60
		Valid.	-0.52	1.49	-	1.04	-	0.80	-0.10	0.90	0.85	-0.04
	E4	Calib.	-0.31	2.35	5.96	1.53	4.56	0.63	0.85	0.95	0.92	1.12
		Valid.	-0.13	3.16	-	2.30	-	0.31	0.39	0.66	0.65	0.62
Bauru-Caiuá	E3	Calib.	-0.39	1.52	6.66	0.89	5.49	0.63	0.63	0.79	0.85	0.74
		Valid.	-0.36	2.50	-	1.66	-	0.34	-1.48	0.68	0.33	3.37
	E4	Calib.	-0.40	1.91	6.85	1.12	5.38	0.79	0.54	0.89	0.91	0.56
		Valid.	-0.41	1.10	-	1.93	-	0.64	0.24	0.83	0.91	1.72
Parecis	E3	Calib.	-0.56	1.73	9.68	1.20	8.22	0.84	-1.04	0.91	0.87	3.04
		Valid.	-0.59	1.13	-	0.81	-	0.90	-0.52	-0.15	0.26	0.34
	E4	Calib.	-0.44	2.41	9.12	1.36	6.98	0.71	0.80	0.84	0.87	0.95
		Valid.	-0.41	2.97	-	2.36	-	0.28	0.41	0.61	0.89	0.58
Urucuia	E3	Calib.	-0.15	0.61	2.81	0.37	2.12	0.78	0.78	0.88	0.86	1.19
		Valid.	-0.41	1.59	-	0.95	-	0.12	0.62	0.90	0.91	0.64
	E4	Calib.	-0.17	0.64	3.25	0.36	2.44	0.74	0.84	0.87	0.92	1.06
		Valid.	-0.10	1.43	-	1.01	-	0.31	0.11	0.35	0.73	0.72
Cabeças	E3	Calib.	-0.22	0.21	5.54	0.12	4.43	0.92	0.82	0.83	1.01	1.05
		Valid.	-0.46	0.88	-	0.59	-	0.32	-5.75	0.36	1.08	-5.08
	E4	Calib.	-0.04	0.41	5.54	0.34	4.27	0.81	0.86	0.90	0.91	0.95
		Valid.	-0.29	0.16	-	0.12	-	0.91	0.89	0.95	0.94	0.91
Alter do Chão	E4	Calib.	-0.02	3.11	9.75	2.36	7.6	0.82	0.35	0.91	0.95	0.36
		Valid.	0.18	3	-	1.82	-	0.87	0.57	0.20	0.10	0.23
Araripe	E4	Calib.	-0.20	0.27	2.16	0.16	1.52	0.92	0.86	0.96	0.99	1.12
		Valid.	-0.30	2.12	-	1.71	-	0.11	-0.11	0.16	0.56	0.76
Içá	E4	Calib.	-0.04	1.11	4.24	0.68	3.13	0.91	-0.15	0.88	0.78	-0.17
		Valid.	-0.01	4.93	-	3.73	-	0.94	0.54	0.98	0.97	1.45