


## Modeling the dissolved oxygen response to phosphorus inputs in Lake Spokane: the fallacy of using complex over-parameterized models as the basis for TMDL decisions

Michael T. Brett & George B. Arhonditsis


To cite this article: Michael T. Brett & George B. Arhonditsis (2016) Modeling the dissolved oxygen response to phosphorus inputs in Lake Spokane: the fallacy of using complex over-parameterized models as the basis for TMDL decisions, Lake and Reservoir Management, 32:3, 280-287, DOI: [10.1080/10402381.2016.1211911](https://doi.org/10.1080/10402381.2016.1211911)

To link to this article: <http://dx.doi.org/10.1080/10402381.2016.1211911>

 Published online: 05 Aug 2016.

 Submit your article to this journal [↗](#)

 Article views: 4

 View related articles [↗](#)

 View Crossmark data [↗](#)

RESPONSE

## Modeling the dissolved oxygen response to phosphorus inputs in Lake Spokane: the fallacy of using complex over-parameterized models as the basis for TMDL decisions

Michael T. Brett<sup>a</sup> and George B. Arhonditsis<sup>b</sup>

<sup>a</sup>Department of Civil and Environmental Engineering, University of Washington, Seattle, WA; <sup>b</sup>Ecological Modeling Laboratory, Department of Physical & Environmental Sciences, University of Toronto, Toronto, Ontario, Canada

We thank Wells and Berger (2016) for sharing their perspective on our analysis of the modeled response of Lake Spokane hypolimnetic dissolved oxygen to phosphorus inputs. Our reply to Wells and Berger's (2016) comment will focus on 6 key points: (1) model sensitivity, (2) model calibration, (3) cryptic phosphorus, (4) sediment oxygen demand, (5) optimal model complexity and future data collection efforts, and (6) the importance of establishing proper criteria for improving the peer-review process of modeling studies.

### Model sensitivity

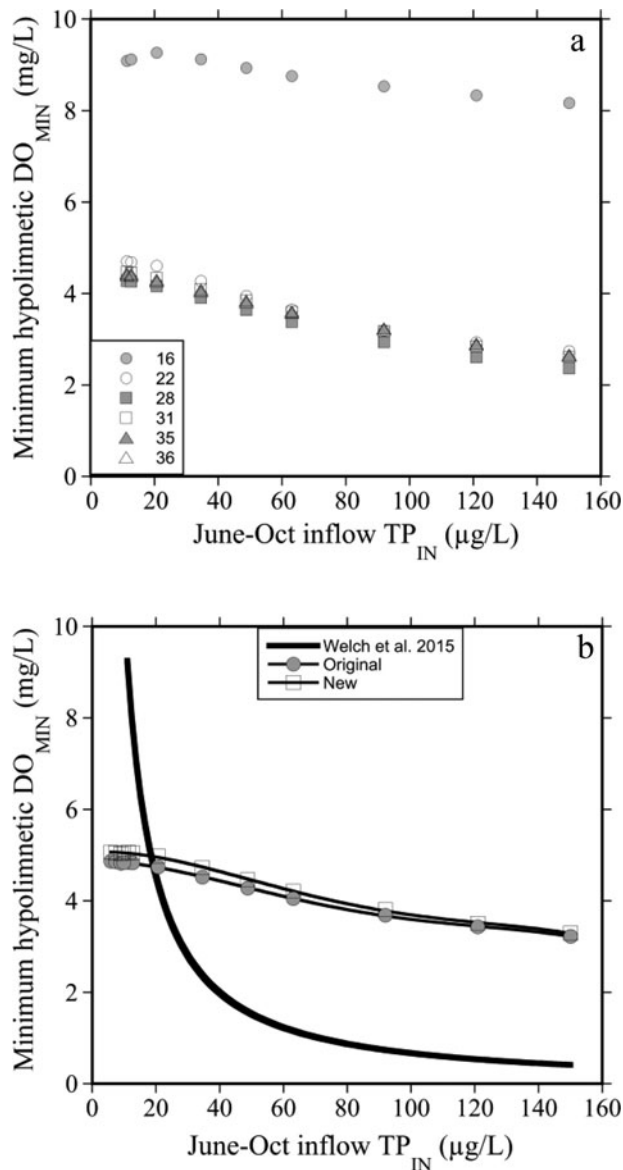
The most important conclusion from our original analysis (Brett et al. 2016) was that the modeled representation of Lake Spokane is relatively insensitive to phosphorus inputs, whereas field data collected from this reservoir suggests it is in fact very sensitive to phosphorus inputs. For example, water quality data compiled by Welch et al. (2015) showed that during its hypereutrophic phase the input phosphorus concentration ( $TP_{IN}$ ) to Lake Spokane averaged  $86 \pm 37$  (SD)  $\mu\text{g/L}$  with the corresponding minimum hypolimnetic oxygen concentration ( $DO_{MIN}$ ) averaging  $1.4 \pm 1.3$  mg/L. Welch et al. (2015) also reported that currently (2010–2014) these values average  $14 \pm 3$   $\mu\text{g/L}$  and  $6.5 \pm 0.8$  mg/L, respectively. We reported that the Lake Spokane water quality (WQ) model's  $DO_{MIN}$  response for similar  $TP_{IN}$  concentrations was much less pronounced, with hypereutrophic and contemporary  $DO_{MIN}$  averaging  $3.8 \pm 0.4$  and  $4.7 \pm 0.04$  mg/L, respectively. We also reported that the WQ model had

a structural DO deficit (saturated DO –  $DO_{MIN}$ ) of 5.3 mg/L when  $TP_{IN}$  was set to zero.

Wells and Berger (2016) suggested our analysis of the Lake Spokane WQ model was biased because we focused on a lake region where (and a time of year when) the lowest DO concentrations were expected. It is true that our analyses focused on the deepest part of Lake Spokane nearest to the dam outlet (i.e., Segment 36) during late August. The location and date were chosen because the water quality standards for the State of Washington indicate DO concentrations shall “not fall below the [one-day minimum DO] criteria . . . at a probability frequency of more than once every ten years.” Therefore, management actions usually target the locations and times when water quality standards are most likely to be violated.

We reanalyzed all of the model outputs from the SPR6 files of the Spokane WQ model to see if our conclusions were biased by our decision to focus on Segment 36. The SPR6 output file reports DO profiles for 6 stations where lake monitoring data are available and with depths of  $\geq 15$  m (i.e., Segments 16, 22, 28, 31, 35, and 36). We reanalyzed the statistical relationship between  $TP_{IN}$  and  $DO_{MIN}$  for each of these stations, separately. We also compiled an overall relationship between these variables by depth weighting the  $DO_{MIN}$  data within and between stations and linearly interpolating these results for all potential Segments between 16 and 36 (i.e., Segments 16, 17, 18 . . . 36). This reanalysis showed the  $DO_{MIN}$  concentrations were considerably higher at model Segment 16 and remarkably similar at Segments 22, 28, 31, 35,

and 36 (Fig. 1a) and that the overall depth/volume and spatially weighted response between  $TP_{IN}$  and  $DO_{MIN}$  was (after correcting for cryptic phosphorus) within 0.2 mg/L of the values that we originally reported (Fig. 1b). These new results indicate the Lake Spokane WQ model predicts hypereutrophic and contemporary  $DO_{MIN}$  averaged  $4.0 \pm 0.4$  and  $5.0 \pm 0.04$  mg/L, respectively, and the WQ model has a structural DO deficit of 5.1 mg/L. Thus, our key conclusion remains



**Figure 1.** The modeled and observed relationship between input total phosphorus concentrations ( $TP_{IN}$ ) and the minimum hypolimnetic dissolved oxygen concentration ( $DO_{MIN}$ ) in Lake Spokane. (a) Volume-weighted responses for model segments 16, 22, 28, 31, 35, and 36. (b) Observed response of Lake Spokane as indicated by field data (i.e., Welch et al. 2015), the modeled response originally reported by Brett et al. (2016), and the revised depth and spatially weighted response calculated in this study.

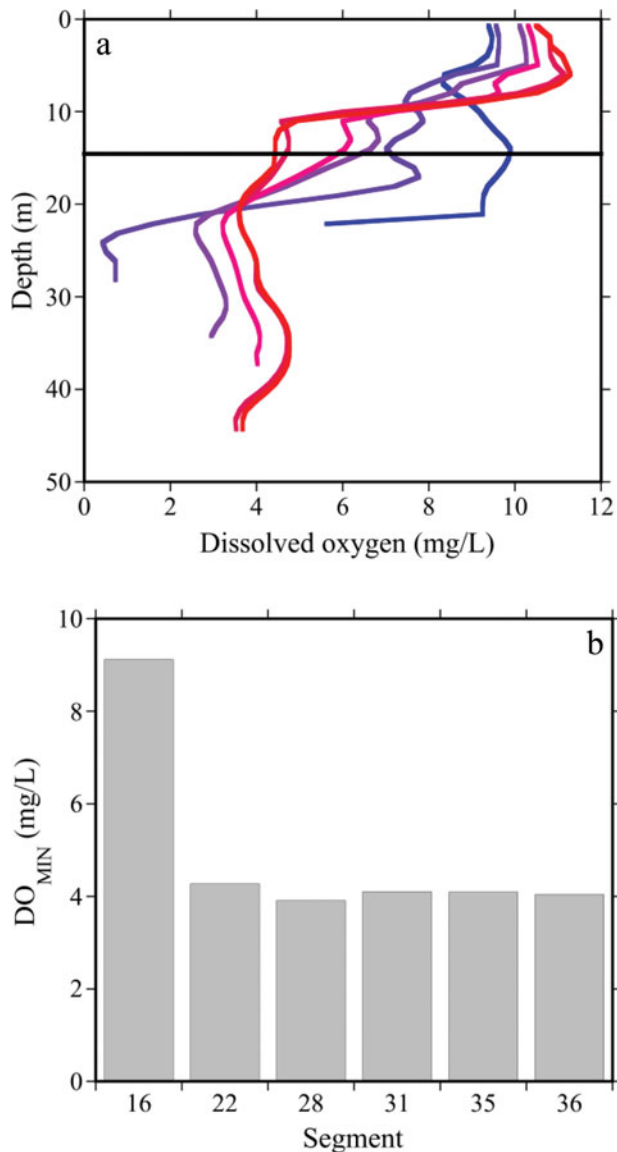
unchanged: the Lake Spokane field data showed a 5.1 mg/L increase in the minimum hypolimnetic DO concentration when the reservoir shifted from hypereutrophy to meso-oligotrophy (Welch et al. 2015), whereas for the same  $TP_{IN}$  concentrations, the model representation of Lake Spokane showed only a 1.0 mg/L improvement in  $DO_{MIN}$ .

Because Wells and Berger (2016) claimed that the modeled  $DO_{MIN}$  values are strongly dependent on the reservoir segment considered, we also explored the spatial trends in the SPR6 DO profiles during late August for the effluent  $TP = 500 \mu\text{g/L}$  scenario (which roughly corresponds to current effluent concentrations). This comparison showed that although vertical profiles varied from one model segment to another (Fig. 2a), the  $DO_{MIN}$  values for Segments 22 through 36 were remarkably similar (Fig. 2b).

Wells and Berger (2016) also claimed the  $DO_{MIN}$  trends we reported were temporally transient, but closer inspection of the SPR6 output files did not support this conjecture. Instead these data showed the model predicted an 8-week period of low hypolimnetic oxygen concentrations extending from late July to late September at model Segments 22 to 36. Because Wells and Berger (2016) suggested a strong spatial gradient in the  $DO_{MIN}$  values (figure 4 in Wells and Berger 2016) and our results suggested very similar values for model Segments 22 through 36 (our Fig. 1a and 2b), we also explored this point for the total maximum daily load (TMDL) scenario used by Wells and Berger (2016). In Wells and Berger's (2016) figure 4, the  $DO_{MIN}$  value for > 15 m at Segment 22 in the TMDL scenario was 7.7 mg/L, similar to the model output for the volume-weighted depths > 8 m (i.e., 7.3 mg/L) but very different from our result for the same segment and scenario in late August (i.e., 4.7 mg/L). We plotted the DO vertical profile for Segment 22 for depths > 15 m for the same scenario to clarify this discrepancy (Fig. 3). This profile showed only one depth with a DO concentration > 7.7 mg/L, and most depths had DO concentrations well below that value. These data indicate that a volume-weighted > 15 m  $DO_{MIN}$  concentration of 7.7 mg/L for this segment and scenario during late August is implausible.

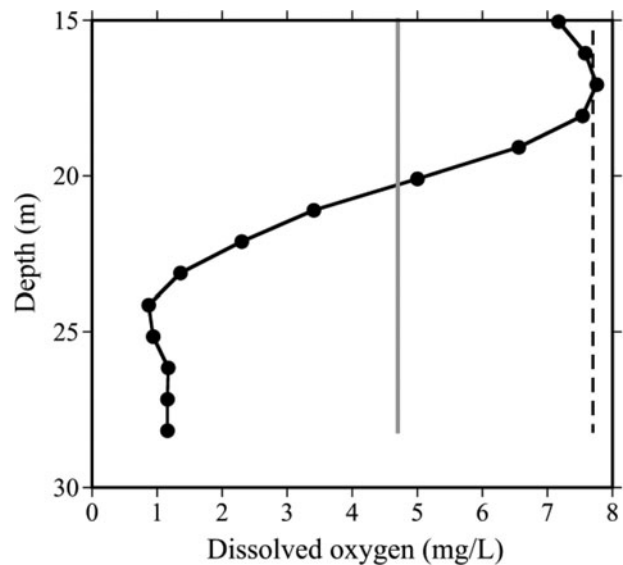
### Model calibration

Wells and Berger (2016) point out that the lack of consideration of the model calibration against data from



**Figure 2.** (a) Modeled dissolved oxygen depth profiles and (b) volume-weighted  $DO_{MIN}$  values for Segments 16 through 36 in late August for the effluent  $TP = 500 \mu\text{g/L}$  scenario. In panel (a), dark blue represents Segment 16, and bright red represents Segment 36, with the intermediate segments having intermediate colors.

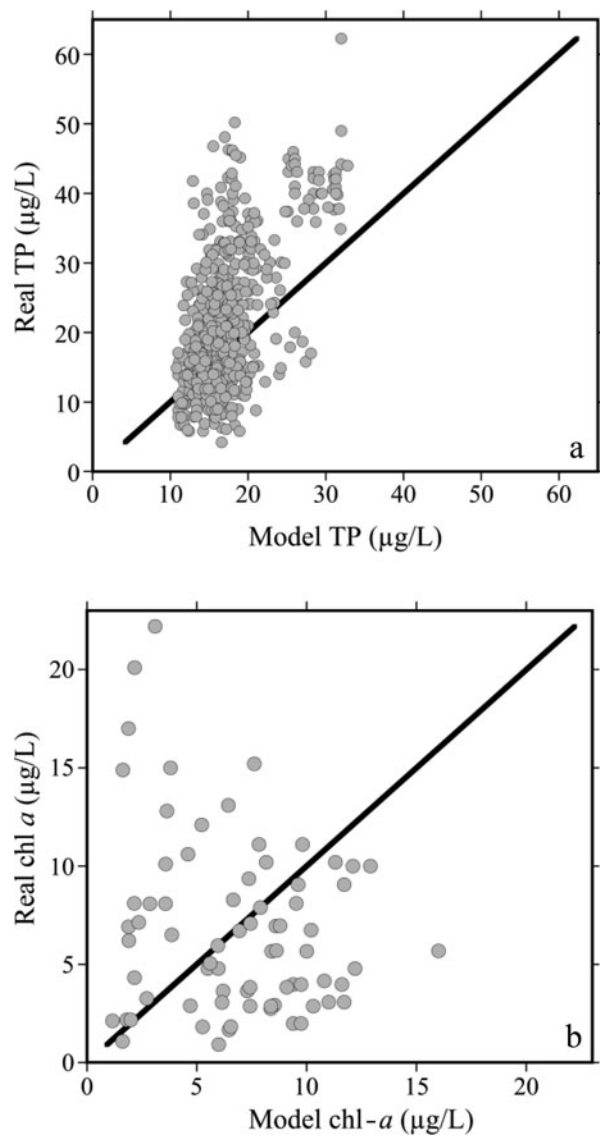
1991 and 2000 was an oversight of our recent analyses. Because we did not have these raw data, we digitized the TP and chlorophyll *a* (Chl-*a*) data plotted in Berger et al. (2002) to test how well the model was calibrated to the TP and Chl-*a* field data for those years. We did not quantify the calibration fits for the temperature and DO data from Berger et al. (2002) because visual inspection suggested those data had adequate fits, as previously noted by Brett et al. (2016) for the 2001 data. The calibration fit for the 1991 and 2000 TP data was slightly better than the fit reported for 2001 data in Brett et al. (2016); the Nash–Sutcliffe model efficiency index was



**Figure 3.** The dissolved oxygen depth profile for Segment 22 for the TMDL scenario in late August (filled circles). This panel also shows the volume-weighted  $DO_{MIN}$  concentration for  $>15 \text{ m}$  that Wells and Berger (2016) reported for Segment 22 for the TMDL scenario in their figure 4 (dashed vertical line) and the volume-weighted  $DO_{MIN}$  concentration that we calculated for this profile (gray vertical line).

–0.10 for the 2001 and 0.09 for the 1991 and 2000 data, respectively (Table 1). When the observed TP concentrations were  $\leq 11.0 \mu\text{g/L}$ , the model always overestimated TP (mean error [ME] =  $6.5 \mu\text{g/L}$ ), and when the observed TP concentrations were  $\geq 23 \mu\text{g/L}$ , the model always underestimated TP (ME =  $-12.7 \mu\text{g/L}$ ) (Fig. 4a).

The calibration fit for 2000 Chl-*a* data was much worse than the fit reported in Brett et al. (2016); the model efficiency index (MEI) was 0.14 for the 2001 and –0.85 for the 2000 data, respectively (Table 1). The highly negative MEI value for the 2000 Chl-*a* data was because the field and model Chl-*a* data were weakly negatively correlated (Fig. 4b), and the model error sum of squares was 85% larger than the total sum of squares for the observed Chl-*a* data. This finding indicates that the model predictions were much worse than a prediction based on the average of the observed values. These results and visual inspection of the calibration data from 1991 and 2000 (Fig. 4) strongly reinforce the conclusion in Brett et al. (2016) that the Spokane WQ model is poorly calibrated to the mechanistic basis of the Spokane basin



**Figure 4.** Bivariate plots of the model and field data for (a) total phosphorus (TP) and (b) chlorophyll *a* (Chl-*a*) reported in the Berger et al. (2002) Lake Spokane WQ model calibration report. The TP data were collected during 1991 and 2000, and the Chl-*a* data were collected in 2000.

TMDL (i.e., DO dependence on  $TP_{IN}$  via phytoplankton biomass production).

Wells and Berger (2016) also questioned our use of 2001 baseline conditions. We used 2001 conditions

for our analyses because this hydraulic year has been selected as the baseline condition for TMDL scenario testing (Moore and Ross 2010). According to the 2010 TMDL report for the Spokane Basin “the [2001] low river flow period is expected to be the most critical period for pollutant loading effects in the river and Lake Spokane due to less dilution of nutrient concentrations and a longer retention time, both of which can exacerbate dissolved oxygen shortages” (Moore and Ross 2010). The Welch et al. (2015) results show that  $DO_{MIN}$  is correlated with hydraulic inflows, and therefore the model should be less responsive to phosphorus inputs when inflows are greater. As far as we are aware, however, this prediction has not been assessed with a proper model sensitivity analysis.

Our analysis used the exact same model inputs as the official TMDL scenario (Berger et al. 2009, Moore and Ross 2010) for all variables except  $TP_{IN}$ , which was systematically varied over a wide range of concentrations in a classic one-step-at-a-time sensitivity analysis to test the model  $DO_{MIN}$  response to a range of nutrient inputs. If our analysis is invalid because it depends on the 2001 hydrologic year, then it would be equally true that the entire TMDL decision making process for the Spokane Basin is invalid because it is also based on the 2001 hydrology conditions. Our  $TP_{IN}$  sensitivity analysis showed the Spokane WQ model had a small response to phosphorus inputs. Had we picked less extreme conditions (such as higher inflows), the model’s responses to phosphorus inputs likely would have been even smaller. Further, the notion that the model  $DO_{MIN}$  values should respond to different inflows is a logical expectation based on how most real reservoirs would respond, but it is not currently known whether the Spokane WQ model responds as it should to flow. Because we found the model was rather insensitive to  $TP_{IN}$ , it is also conceivable that the modeled  $DO_{MIN}$  is insensitive to the average inflow for the range of conditions typically encountered in Lake Spokane.

**Table 1.** The results of the model calibration analysis based on the observations reported in Berger et al. (2002). The Nash–Sutcliffe model efficiency index ( $r^2$ ), nonparametric coefficient of determination (N-P  $r^2$ ), the root mean square error (RMSE), mean error (ME), mean absolute error (MAE), and relative error (RE) are reported. The total phosphorus results are based on field data collected in 1991 and 2000, and the chlorophyll *a* results are based on data collected in 2000.

	Count	Mean $\pm$ SD	Units	MEI ( $r^2$ )	N-P $r^2$	RMSE	ME	MAE	RE
Total Phosphorus	619	21.6 $\pm$ 9.6	$\mu\text{g/L}$	0.09	0.22	9.2	– 4.7	7.0	0.32
Chlorophyll <i>a</i>	74	6.7 $\pm$ 4.5	$\text{mg/L}$	– 0.85	0.00	6.1	0.1	4.6	0.69

## Cryptic phosphorus

We reported that when all phosphorus inputs to the Spokane River were set to zero (including the wastewater treatment plants, the Lake Coeur d'Alene outlet, Hangman Creek, Kaiser Aluminum, stormwater, combined sewer overflows, groundwater, and various smaller discharges), the  $TP_{IN}$  concentration averaged 1.1  $\mu\text{g/L}$  (Brett et al. 2016). Further, when we set all TP inputs to Lake Spokane to zero (including the Little Spokane River and direct groundwater inputs to the reservoir), epilimnetic TP concentrations in Lake Spokane averaged 3.7  $\mu\text{g/L}$  (Brett et al. 2016). We referred to this pool as cryptic phosphorus because it had no obvious external source. Wells and Berger (2016) pointed out that what we called cryptic phosphorus was most likely the combination of phosphorus bound in algal biomass and labile and recalcitrant dissolved and particulate organic matter (i.e., LDOM-P, RDOM-P, LPOM-P, and RPOM-P, respectively). The former constituent of the “correct”  $TP_{IN}$  was simulated as the product of the (poorly reproduced) phytoplankton biomass variability with a fixed stoichiometry, a practice frequently challenged in the modeling literature (Zhao et al. 2008). The latter 4 phosphorus pools were only vaguely mentioned in the documentation for the Spokane WQ model. For example, Berger et al. (2009) stated “since organic matter originating from point sources and tributaries was modeled with CBOD [carbonaceous biological oxygen demand] compartments, the labile dissolved organic matter (LDOM), refractory dissolved organic matter (RDOM), labile particulate organic matter (LPOM), and refractory particulate organic matter (RPOM) compartments only simulated the by-products of phytoplankton and periphyton decay.” Berger et al. (2009) also stated “since organic matter was accounted for in the CBOD compartment, constituent concentrations of the fractional organic matter compartments—LDOM, RDOM, LPOM, and RPOM—were set to zero.” Berger et al. (2002) had the same descriptions of these pools.

We checked the magnitude of these additional phosphorus pools in the CWO\_151 output files and found that in the effluent  $TP = 0, 50, 500$  and  $2000 \mu\text{g/L}$  scenarios, the sum of algal-, DOM-, and POM-bound phosphorus averaged 3.9, 4.3, 6.6 and 9.3  $\mu\text{g/L}$ , respectively, during June to October. In these cases, algal-P ranged between 0.8 and 2.8  $\mu\text{g/L}$ , LDOM-P between 2.0 and 5.1  $\mu\text{g/L}$ , LPOM-P between 0.2 and 0.6  $\mu\text{g/L}$ ,

and RDOM-P and RPOM-P were nearly constant at  $0.9 \pm 0.0$  and  $0.0 \pm 0.0 \mu\text{g/L}$ , respectively. With these data, we were able to correct the  $TP_{IN}$  data for the previously mentioned  $DO_{MIN}$  sensitivity analyses. This correction showed that the discrepancy in the  $TP_{IN}$  calculation had virtually no implications for any of the modeling experiments or points raised by Brett et al. (2016). Accounting for algal- and DOM/POM-bound P only affected the  $TP_{IN}$  values by a few micrograms per liter at the lower range of  $TP_{IN}$ , and in this region  $DO_{MIN}$  was insensitive to phosphorus inputs (Fig. 1b). For example, when using the corrected input data,  $DO_{MIN}$  averaged  $5.04 \pm 0.03 \text{ mg/L}$  for  $TP_{IN}$  values ranging from 6 to 21  $\mu\text{g/L}$  (Fig. 1b).

Interestingly, the sum of phosphate, CBOD-P, algal-P, LDOM-P, RDOM-P, LPOM-P, and RPOM-P in the zero phosphorus loading scenario for June to October equaled 2.8  $\mu\text{g/L}$ , which closely matched the epilimnetic TP concentration in the corresponding Lake Spokane zero loading scenario (Brett et al. 2016). It is still unclear how this phosphorus got into the model representation of the Spokane River because we set all inputs to the river to zero, and, as noted by Berger et al. (2009), LDOM, RDOM, LPOM, and RPOM originating from point sources and tributaries were also set to zero by default. Perhaps, it was a residual effect of the initial concentrations for algal and DOM/POM bound phosphorus (i.e., 3.5  $\mu\text{g/L}$ ), or perhaps the model's representation of the river periphyton retained some P in the system, even in the absence of inputs.

For example, the model predicts a very large periphyton biomass in the river, which sequesters an exceptionally large pool of phosphorus. Therefore, the model representation of periphyton may act as an important phosphorus pool that does not actually exist in the real Spokane River. During the summer months, the model predicts the periphyton has an average biomass of 18.3  $\text{mg/L}$  in the TMDL scenario at river mile 151, and that this periphyton has a stoichiometric equivalent between phosphorus and organic matter of 0.005 (Berger et al. 2002). This biomass equates to a periphyton phosphorus content of 92  $\mu\text{g/L}$ , a pool more than 9 times larger than the average TP pool size in the river water itself (i.e., mean  $TP = 9.8 \mu\text{g/L}$ ) during the summer. The model also predicts periphyton biomass fluctuates by an average of 6% on a daily basis in a sinusoidal pattern, which also equates to a very large flux of riverine phosphorus

into and out of the periphyton biomass. This suggests phosphorus exchange between the periphyton and the overlying river water is a critical aspect of how phosphorus transport in the Spokane River is modeled. Because Welch et al. (1989) quantified periphyton biomass in the lower Spokane River for a wide range of soluble reactive phosphorus (SRP) concentrations, we can also ask whether the periphyton biomass values output in the Spokane WQ model are reasonable. Welch et al. (1989) found that periphyton biomass measured as Chl-*a* averaged  $49 \pm 45$  (SD) mg/m<sup>2</sup> for river water samples with SRP concentrations <10 µg/L. If we use the 65 to 1 conversion between periphyton biomass and Chl-*a* from Berger et al. (2002), then this equates to an overall organic matter biomass of  $3200 \pm 2900$  mg/m<sup>2</sup>. If we further assume the lower Spokane River has a mean depth of 1.5 m, this equates to a periphyton volume concentration of  $2.1 \pm 2.0$  mg/L. This value suggests the periphyton biomass values output by the Spokane WQ model may be 8 to 9 times larger than that observed in the lower Spokane River for equivalent riverine phosphorus concentrations.

### Sediment oxygen demand

In their critique, Wells and Berger (2016) stated the Spokane WQ “model was misapplied [by Brett et al. (2016)] by assuming that the model’s zero order sediment oxygen demand [SOD] was not dependent on phosphorus inflows.” This point implies that, in general, the Spokane WQ model dynamically couples SOD to phosphorus concentrations; however, a close inspection of past modeling reports does not support this. For example, Berger et al. (2002) reported that during the 1991 and 2000 calibrations “zero order sediment oxygen demand rates were set at 0.6 g/m<sup>2</sup>/d for [all] Long Lake Reservoir model segments,” despite Lake Spokane TP concentrations averaging 21% higher during 1991. During the 2001 calibration, Wells and Berger (2016) stated the “SOD of model segments was set to values between 0.1 to 0.6 g/m<sup>2</sup>/d, with higher SOD in model segments closer to the dam.” In this case, substantially lower SOD values were assumed for some reservoir segments during 2001 than during 2000, even though the mean reservoir TP values were similar both years. For the TMDL scenario, Wells et al. stated, “SOD set to improved level = 0.25 g DO/m<sup>2</sup>/d” (PSU 2010). In aggregate, it appears the process Wells and colleagues used for assigning SOD values in the Spokane WQ

model was not systematically based on phosphorus inputs or concentrations. Furthermore, as far as we are aware, none of these assumed SOD values were based on actual field SOD determinations from the particular reservoir segments and years modeled. As noted by Rucinski et al. (2014) and reiterated by Brett et al. (2016), the SOD rates used in lake or reservoir WQ models should also be dynamically coupled to nutrient inputs and phytoplankton production.

In the future, it is important that a systematic and repeatable algorithm be developed for Lake Spokane that couples the SOD rates used in the model to the trophic state of the reservoir. It is also essential that this algorithm be properly documented and undergo thorough peer review. Until these augmentations are in place, we concur with the CE-QUAL-W2 manual (page 14; Cole and Wells 2015) that the lack of “. . . a sediment compartment that models kinetics in the sediment and at the sediment–water interface, i.e., a complete sediment diagenesis model. . . places a limitation on long-term predictive capabilities of the water quality portion of the model. . .”

### Optimal model complexity and future data collection

The Spokane WQ model is a characteristic case of a construct not commensurate with the available data from the system. Indicative of the profound underdetermination problem is the TP formula provided by Wells and Berger (2016), in which available information exists only for 1 or 2 state variables among the 9 major constituents simulated. An example of the problematic decoupling between model complexity and empirical evidence was the inclusion of 3 algal groups and the questionable characterization of their functional properties, reflective of the tendency in aquatic biogeochemical modeling to “run before we can walk” (Anderson 2005, Shimoda and Arhonditsis 2016). None of the reports published by Wells and Berger over the past 15 years have attempted to shed light on the potential implications of the unconstrained state variables for the predictive statements drawn from their modeling exercise. Although the increase of the articulation level is certainly an effective means for improving our models, we should not ignore that the increasing complexity also reduces our ability to properly constrain the model parameters from observations; that is, the number of

parameters that must be specified from the data is approximately proportional to the square of the number of compartments (Denman 2003). Thus, rather than first establishing the fundamental relationships among nutrient loading, in-stream conditions, and a generic phytoplankton compartment (parameterized against the available Chl-*a* data), they simulated 3 ambiguously defined algal groups, as explicitly recognized by Wells and Berger (2016), with little consideration of the state of knowledge of phytoplankton functional group ecology and their associated mathematical representation (Shimoda and Arhonditsis 2016).

Brett et al. (2016) argued that the Spokane WQ model offers a “unique opportunity to test the ability of mechanistic models to simulate natural conditions because it has a detailed observational record of water quality conditions.” With the present commentary, we also wish to emphasize that the identified problems with the determinacy state of the existing model can be used to guide data collection efforts. For example, Wells and Berger (2016) pinpointed the relationship between upstream phosphorus loading and SOD as a critical causal linkage for projecting future system responses. Detailed knowledge of the processes occurring in the top few centimeters of the sediment is essential for assessing water quality, understanding the manifestation of hypoxia, and managing surface waters. Field, experimental, and modeling work should be designed to elucidate the mechanisms of phosphorus mobilization and oxygen utilization in the sediments and to identify process controls under a variety of conditions.

## Conclusions

A reanalysis of the hypolimnetic DO concentrations predicted by the Spokane WQ model showed that a large spatial region (i.e., Segments 22 to 36) in the model representation of this reservoir has little sensitivity to phosphorus inputs and has a large structural oxygen deficit. Analyses of additional model calibration data collected in 1991 and 2000 showed that the TP dynamics in the model are poorly correlated to the field data, and the modeled Chl-*a* data are inversely correlated with the field data. This finding reinforces the conclusions of our recent study that showed the Spokane WQ model is not properly calibrated to the mechanistic basis of the Spokane Basin dissolved oxygen TMDL (Brett et al. 2016). It is critical that the Spokane WQ model be recalibrated and that

contemporary field data for Lake Spokane be given much more emphasis in the TMDL decision-making process.

On a final note, we could not agree more with the Wells and Berger's (2016) plea for improving the peer review process of modeling studies. In several recent papers (Arhonditsis and Brett 2004, Arhonditsis 2009), we have emphatically underscored the importance of establishing a systematic methodological protocol (detailed sensitivity analysis, rigorous uncertainty analysis, realistic parameter specification that capitalizes on our knowledge of the ecosystem functioning) for aquatic biogeochemical model development along with performance criteria (explicit specification of what constitutes an “acceptable model error” given the questions being addressed by the model) widely adopted by the modeling community. Even if journals cannot enforce the submission of all the material required to reconstruct mathematical models (Flynn 2005), they should demand that modeling studies meet specific methodological and performance criteria. In our opinion the Spokane WQ model does not meet these criteria, but we sincerely hope that this discussion will lead to improved model-based water quality management in the Spokane Basin.

## References

- Anderson TR. 2005. Plankton functional type modelling: running before we can walk? *J Plankton Res* 27:1073–1081.
- Arhonditsis GB. 2009. Useless arithmetic? Lessons learnt from aquatic biogeochemical modelling. In: Hanrahan G, editor. *Modelling of pollutants in complex environmental systems*. ILM Publications. p. 3–26.
- Arhonditsis GB, Brett MT. 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. Where are we? *Mar Ecol Prog Ser* 271:13–26.
- Berger CJ, Annear RL, Wells SA. 2002. Upper Spokane River model: model calibration. Portland (OR): Portland State University, Department of Civil Engineering. Technical Report EWR-1-02.
- Berger CJ, Wells SA, Wells V. 2009. Spokane river modeling report 2009, model development and calibration check. Portland (OR): Portland State University, Department of Civil Engineering. Technical Report EWR-03-09.
- Brett MT, Ahopelto SK, Brown HK, Brynestad BE, Butcher TW, Coba EE, Curtis CA, Dara JT, Doeden KB, Evans KR, et al. 2016. The modeled and observed response of Lake Spokane hypolimnetic dissolved oxygen concentrations to phosphorus inputs. *Lake and Reserv Manage* 32:243–255.
- Cole TM, Wells SA. 2015. CE-QUAL-W2: a two-dimensional, laterally averaged, hydrodynamic and water quality model,



- version 3.72. Portland (OR): Portland State University, Department of Civil and Environmental Engineering.
- Denman KL. 2003. Modelling planktonic ecosystems: parameterizing complexity. *Prog Oceanogr* 57:429–452.
- Flynn KJ. 2005. Castles built on sand: dysfunctionality in plankton models and the inadequacy of dialogue between biologists and modellers. *J Plankton Res* 27:1205–1210.
- Moore DJ, Ross JD. 2010. Spokane River and Lake Spokane dissolved oxygen total maximum daily load: water quality improvement report. State of Washington, Department of Ecology.
- [PSU] Portland State University. 2010. Spokane River modeling final scenarios report 2010. Portland (OR): Portland State University, Department of Civil Engineering. Technical Report EWR-01-10.
- Rucinski DK, DePinto JV, Scavia D, Beletsky D. 2014. Modeling Lake Erie's hypoxia response to nutrient loads and physical variability. *J Great Lakes Res* 40:151–161.
- Shimoda Y, Arhonditsis GB. 2016. Phytoplankton functional type modelling: running before we can walk? A critical evaluation of the current state of knowledge. *Ecol Model* 320:29–43.
- Welch EB, Brattebo SK, Gibbons HL, Plotnikoff RW. 2015. A dramatic recovery of Lake Spokane water quality following wastewater phosphorus reduction. *Lake Reserv Manage* 31:157–165.
- Welch EB, Horner RR, Patmont CR. 1989. Prediction of nuisance periphytic biomass: a management approach. *Water Res* 23:401–405.
- Wells SA, Berger CJ. 2016. Modeling the response of dissolved oxygen to phosphorus loading in Lake Spokane. *Lake Reserv Manage*. 32:280–287.
- Zhao J, Ramin M, Cheng V, Arhonditsis GB. 2008. Competition patterns among phytoplankton functional groups: how useful are the complex mathematical models? *Acta Oecol* 22:324–344.