Chemometric model development for an online water quality monitor for maturation ponds: A case study

Presented by

C K Hasan[1], Linhua Fan[1], Oliver A.H. Jones[1], Nick Crosbie[2], and Felicity Roddick[1]

[1]RMIT University, City Campus

[2]Melbourne Water

E-mail: Chowdhury.Kamrul.hasan@rmit.edu.au

RMIT UNIVERSITY

Background

- Wastewater Treatment Plants mostly rely on grab sampling for monitoring the key water quality parameters (e.g. COD, DOC, NH_4, NO_3, PO_4, and suspended solids).
- Grab sampling is costly and time consuming and does not represent the episodic variations in influents.

- **Challenge:**
  - Significant challenge for the water utility companies to implement real-time monitoring of these parameters.
  - **UV-vis spectrometer** can overcome this challenge.
  - Lack of **suitable** matrix-specific chemometric models for such parameters.
- **Aims:**
  - Development of suitable chemometric models for an intelligent spectral analyser **(ISA) UV–vis spectrometer.**
  - Test the models to an **onsite ISA UV-Vis spectrometer** for monitoring the effluent from maturation ponds.

Novelty and Research Gaps

- **Partial Least Squares** is the widely used machine learning algorithm to develop chemometric models for the prediction of water quality parameters, e.g. COD and TSS.

- Little has been done on **Multiple Linear Regression** to develop chemometric models for **a battery of parameters**, such as COD, DOC, NH_4-N, NO_3-N, PO_4-P, and suspended solids.

- Demonstration of a real-time and an *in-situ* optical sensing device for monitoring of these water quality parameters at a wastewater treatment pond.

RMIT UNIVERSITY

Next…How did we develop the models?

Chemometric Model Development Procedure

1. Spectra **selection** & **indexing**
2. Combining spectra (JDX file)
3. Merging spectra with lab values (CPF file)
4. Generating chemometric models (e.g. TXT or XML files)
5. Loading the files to a lab-based ISA UV-vis

Lab-based ISA UV–vis for the optimisation of the chemometric models prior to applying them to an online ISA UV-vis.

6. Remotely loading the files to the onsite ISA UV–vis installed at a maturation pond

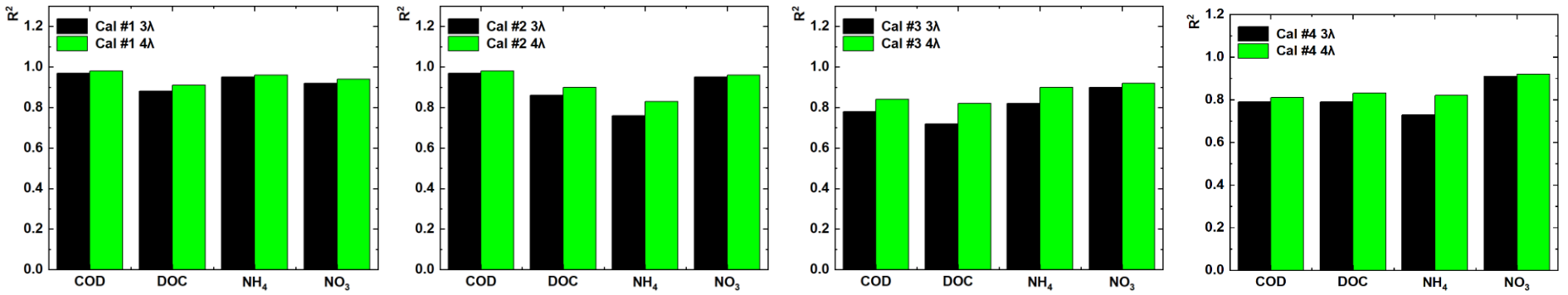
Onsite ISA UV–vis installed at Pond 10 of the maturation pond system.

RMIT UNIVERSITY

Next…mixed-matrix models.

Mixed-matrix Chemometric Models: lab-based

- Developed 4 **mixed-matrix?** models…dataset came from different ponds…
- Validated the models based on higher R^2 values
- …models based on 4-wavelength combinations provided higher R^2 values than those of 3-wavelength combinations…



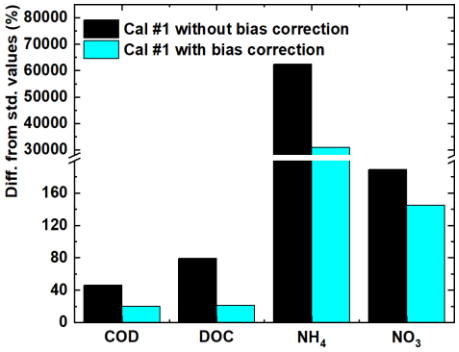
Cal. #1: 60-samples including AP4 and CP10 data, Cal. #2: 52-samples excluding AP4 data, Cal. #3: 44-samples excluding AP4 and CP10 data
Cal. #4: 62-samples (44-samples excluding AP4 & CP10 data sets, but including plant supplied 18-sample for AP10)

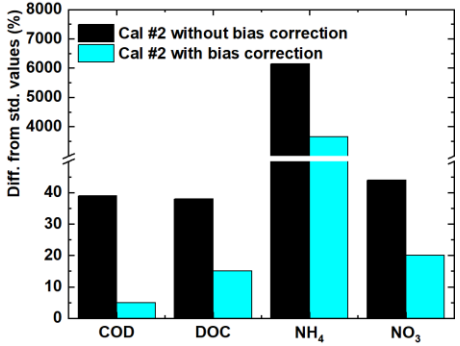
RMIT UNIVERSITY

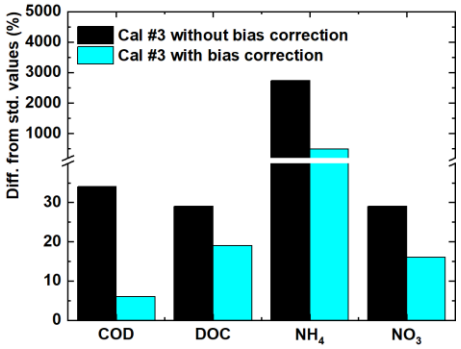
Next…application of mixed-matrix models to a lab-based ISA UV-vis.

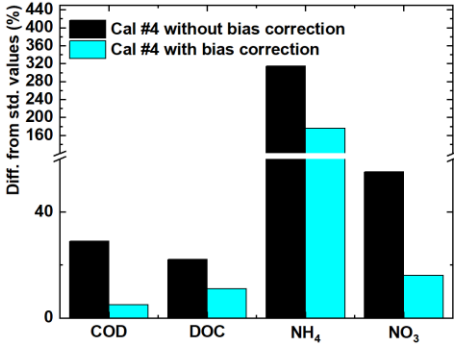
Application of Mixed-matrix Chemometric Models: lab-based

- …the models, without bias correction, **significantly overestimated/underestimated** the concentrations of the parameters, compared to the standard measurements recorded from the samples collected from **the Pond #10.**









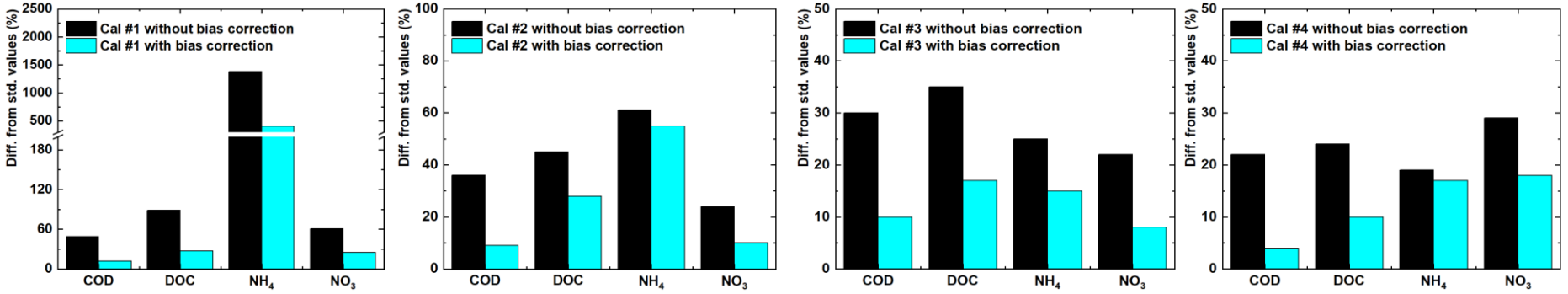
Cal. #1: 60-samples including AP4 and CP10 data, Cal. #2: 52-samples excluding AP4 data, Cal. #3: 44-samples excluding AP4 and CP10 data
Cal. #4: 62-samples (44-samples excluding AP4 & CP10 data sets, but including plant supplied 18-sample for AP10)

RMIT UNIVERSITY

Next…application of mixed-matrix models to pond 5.

Application of Mixed-matrix Chemometric Models: lab-based

- When applied the models to the samples collected from the **pond #5?…**
- **…observed similar trend of overestimation/underestimation** of the parameters…



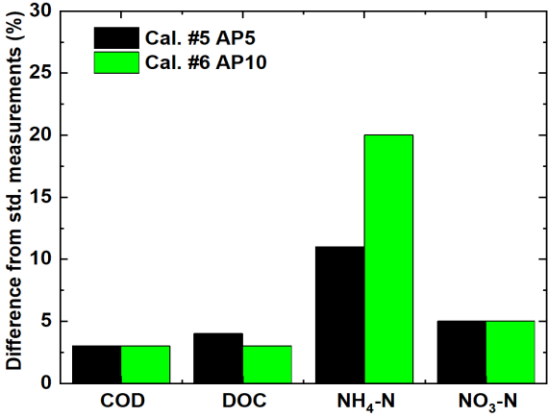
Cal. #1: 60-samples including AP4 and CP10 data, Cal. #2: 52-samples excluding AP4 data, Cal. #3: 44-samples excluding AP4 and CP10 data
Cal. #4: 62-samples (44-samples excluding AP4 & CP10 data sets, but including plant supplied 18-sample for AP10)

RMIT UNIVERSITY

Next…decided to develop matrix-specific models.

Application of Matrix-Specific Chemometric Models: lab-based

- Matrix-specific?...dataset came only from the pond 5 and the pond 10.
- Like mixed-matrix, matrix-specific models based on **4-wavelength** combinations gave higher **R^2** and F-statistic than those of 3-wavelength combinations.
- When we applied the models to **Pond 5** and **Pond 10** samples…
- …the **overestimation** of all the water quality parameters significantly **reduced (3-20%) compared** to the mixed-matrix models.
- The **approach developed** at the lab was then applied to an **onsite ISA UV-vis**.



RMIT UNIVERSITY

Next…onsite ISA UV-vis.

Matrix-specific Chemometric Models for an Onsite ISA UV-Vis

- …chemometric models based on plant supplied 37-sample datasets provided higher values of R^2 and the F-statistic.

Parameters	Wavelength range (nm)	R^2	F value	*R^2	*F value	Best wavelength combinations
COD	250-708	0.85	18.89	**0.97**	**53.59**	256, 358, 368, 376
DOC	226-400	0.73	8.58	**1.00**	**524794.20**	288, 298, 304, 366
NH_4	236-708	0.84	16.68	**0.94**	**25.46**	330, 332, 336, 342
NO_3	226-240	0.99	527.28	**0.99**	**527.28**	226, 228, 234, 236
PO_4	236-708	0.94	60.62	**0.98**	**102.22**	334, 376, 576, 598
SS	236-708	0.73	8.98	**0.94**	**32.23**	270, 468, 514, 516

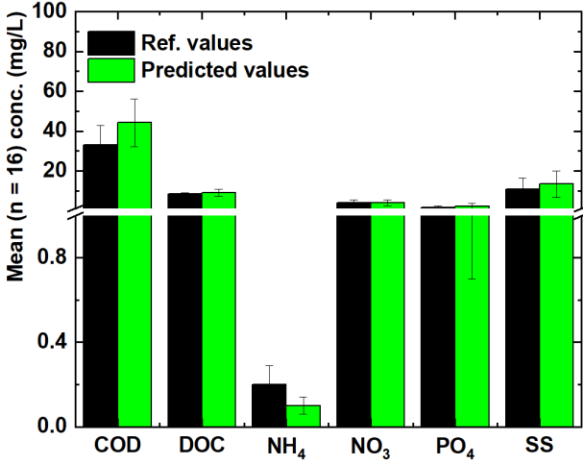
- *Number of samples: 37, no. of wavelengths: 4, H,T, and D outliers: 1, 1, and 1, respectively.*
- **After removing H-outliers of >3.0*

RMIT UNIVERSITY

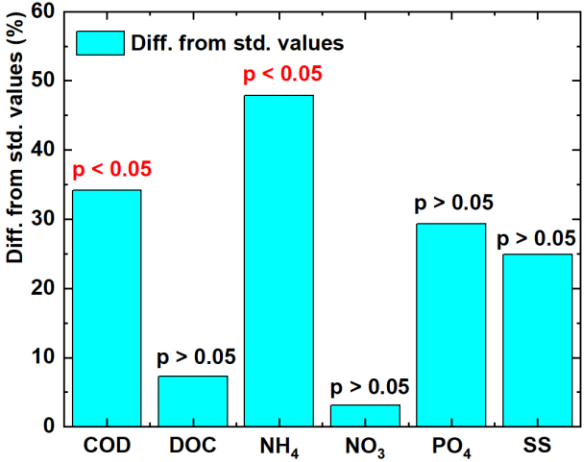
Next…applied to 16 samples from pond 10 and evaluated based on % diff.

Application of the Matrix-specific Chemometric Models: onsite

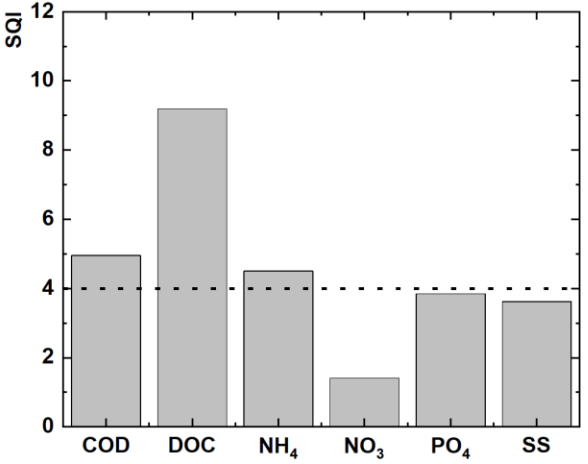
- …the % difference of the predicted values was not significant for **DOC, NO_3, PO_4, and suspended solids**, whereas it was **significant** for **COD** and **NH_4**.



1



2

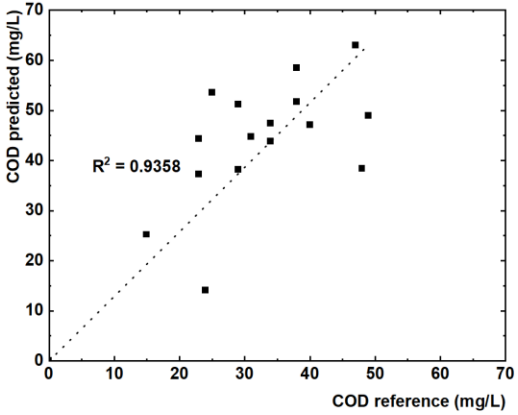


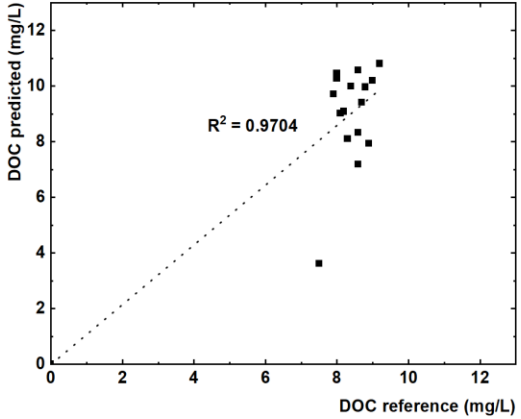
3

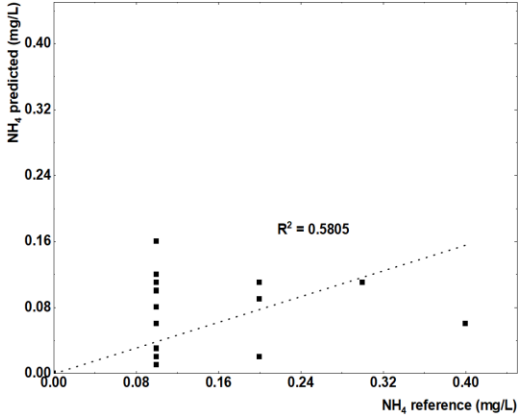
RMIT UNIVERSITY

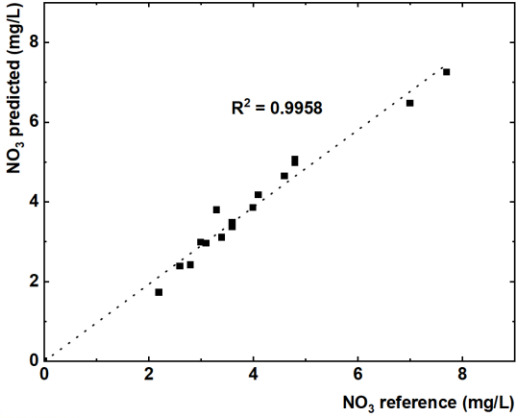
Next… relationship between predicted and ref. values.

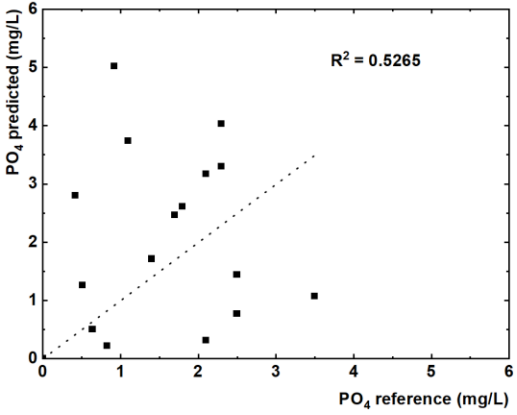
Relationship between Predicted and Reference Values

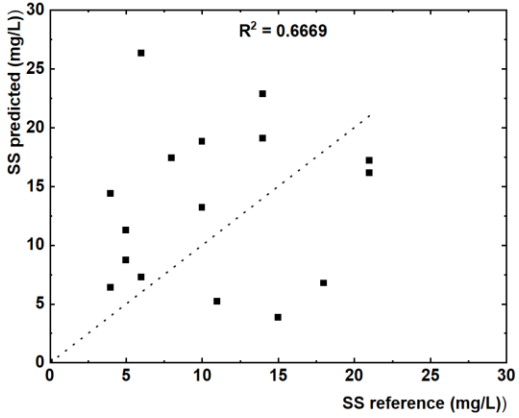












RMIT UNIVERSITY

Next…conclusions

Conclusions and Recommendations

- …improve prediction accuracy of the mixed-matrix models by applying a **bias correction**, but this was **not sufficient to reduce the overestimations** to below 10% for most parameters.

- …the matrix-specific models for the ISA UV-Vis spectrometer have the **potential for *in situ* real-time monitoring** of key water quality parameters **without bias correction**.

- But, the models require **further optimisation**.

RMIT UNIVERSITY