Supporting Information

Spectral Multivariate Calibration without Laboratory Prepared or Determined Reference Analyte Values

Josh Ottaway, Jeremy A. Farrell, and John H. Kalivas*
Department of Chemistry, Idaho State University, Pocatello, Idaho 83209, Untied States

AUTHOR INFORMATION
Corresponding Author

*E-mail: kalijohn@isu.edu.
Fax: 208282-4373

Table of Contents

Subject	Page
PCTR Matlab code	S-2
Figure S-1 (spectral plots of the Inorganic data)	S-4
Table S-1 (concentration design for the Inorganic data set)	S-5
Figure S-2 (spectral plots of the Temperature data)	S-6
Table S-2 (concentration design for the Temperature data set)	S-7

PCTR Matlab Code

```
function out=PCTR(XPC,yPC,NA,yNA,tau,lambda,Xval,yval);
%
%Does PCTR using PC analyte and non-analyte samples
%NOTE: the following code is not optimal and could be speeded up. It is
%written for someone with little knowledge of Matlab and is a brute force
%approach using the equation (10) in the paper
%
%CODE REQUIRES Matlab regstats for obtaining R2, slope, and intercept for
%plotting predicted val sample values against val reference values in yval
%
%INPUT:
%XPC - row vector of pure component analyte spectrum measured at
%        concentration yPC (1 x number of wavelengths)
%yPC - concentration of pure component analyte sample (if pure substance,
%        then yPC = 1
%NA - matrix of non-analyte spectra, if using constant analyte samples, then
%       enter mean centered spectra for NA (number of samples x number of
%       wavelengths)
%yNA - column vector of zeros (concentrations for NA)
%tau - vector of tuning parameter values for the Identity matrix
%lambda - vector of tuning parameter values for NA
%Xval - matrix of validation spectra (number of samples x number of
%        wavelengths)
%yval - column vector of concentrations for Xval
%
%OUTPUT:
%Saved in the out file are all the data that was inputted along with RMSE
%values for the pure component sample, NA, and the val samples, all the
%regression vectors, the respective norms, and the R2, slope, and intercept
%values for the val samples as noted above
%
[ntau,dummy]=size(tau);
[nlambda,dummy]=size(lambda);
[nNA,nwave]=size(NA);
[nval,dummy]=size(Xval);
I=eye(nwave,nwave);
%
%Compute bhats, the model vectors using equation (10) in the paper
for i=1:ntau
   for j=1:nlambda
      [i,j]
      minv=(XPC'*XPC)+((tau(i)^2)*I) + (lambda(j)^2*(NA'*NA));
      %Check to make sure tuning parameter tau is big enough to form a full
      %rank matrix to allow computing the inverse
```

```
        k=rank(minv);
        if k==nwave
            bhat(:,i,j)=inv(minv)*XPC'*yPC;
            nbhat(i,j)=norm(bhat(:,i,j));
            yhat=XPC*bhat(:,i,j);
            rmsec(i,j)=abs(yhat-yPC);
            yhat=NA*bhat(:,i,j);
            rmseNA(i,j)=sqrt((sum((yhat-yNA).^2)/nNA));
            yhat=Xval*bhat(:,i,j);
            rmsev(i,j)=sqrt((sum((yhat-yval).^2)/nval));
            stats=regstats(yhat,yval);
            R2val(i,j)=stats.rsquare;
            interceptval(i,j)=stats.beta(1);
            slopeval(i,j)=stats.beta(2);
        else
            bhat(1:nwave,i,j)=nan;
            nbhat(i,j)=nan;
            rmsec(i,j)=nan;
            rmseNA(i,j)=nan;
            rmsev(i,j)=nan;
            R2val(i,j)=nan;
            interceptval(i,j)=nan;
            slopeval(i,j)=nan;
        end
    end
end
%Create output
out.bhat=bhat;
out.nbhat=nbhat;
out.rmsec=rmsec;
out.rmsen=rmseNA;
out.rmsev=rmsev;
out.R2val=R2val;
out.interceptval=interceptval;
out.slopeval=slopeval;
out.tau=tau;
out.lambda=lambda;
out.XPC=XPC;
out.yPC=yPC;
out.NA=NA;
out.yNA=yNA;
out.Xval=Xval;
out.yval=yval;
```

Spectral Plots of Inorganic Data

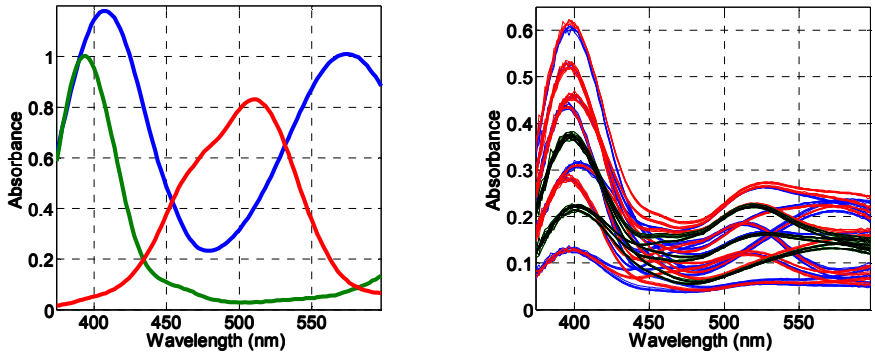

Fig. S-1. Pure component spectra on the right (blue = Cr, green = Ni, and red = Co) and on the right are the RR calibration (blue), validation (red) and constant analyte samples for the non-analyte matrix **N** (black).

Table S-1
Inorganic Concentration Design (Molarity).

Sample type	Cr	Ni	Co
Cal	0.00306	0.0157	0.00688
Cal	0.00306	0.0157	0.0344
Cal	0.00306	0.0471	0.0206
Cal	0.00306	0.0786	0.00688
Cal	0.00306	0.0786	0.0344
Cal	0.00917	0.0157	0.0206
Cal	0.0153	0.0157	0.00688
Cal	0.0153	0.0157	0.0344
Cal	0.0153	0.0471	0.0206
Cal	0.0153	0.0786	0.00688
Non-analyte[a]	0.00917	0.0157	0.00688
Non-analyte	0.00917	0.0157	0.0206
Non-analyte	0.00917	0.0157	0.0344
Non-analyte	0.00917	0.0471	0.00688
Non-analyte	0.00917	0.0471	0.0206
Non-analyte	0.00917	0.0471	0.0344
PC Cr	0.0719	0	0
PC Ni	0	0.0949	0
PC Co	0	0	0.0952
Val	0.00306	0.0157	0.0206
Val	0.00306	0.0471	0.00688
Val	0.00306	0.0471	0.0344
Val	0.00306	0.0786	0.0206
Val	0.00917	0.0786	0.00688
Val	0.00917	0.0786	0.0344
Val	0.0153	0.0157	0.0206
Val	0.0153	0.0471	0.00688
Val	0.0153	0.0471	0.0344
Val	0.0153	0.0786	0.0206

[a]All non-analyte samples are constant analyte and hence, mean centered before being placed in **N**.

Spectral Plots of Temperature Data

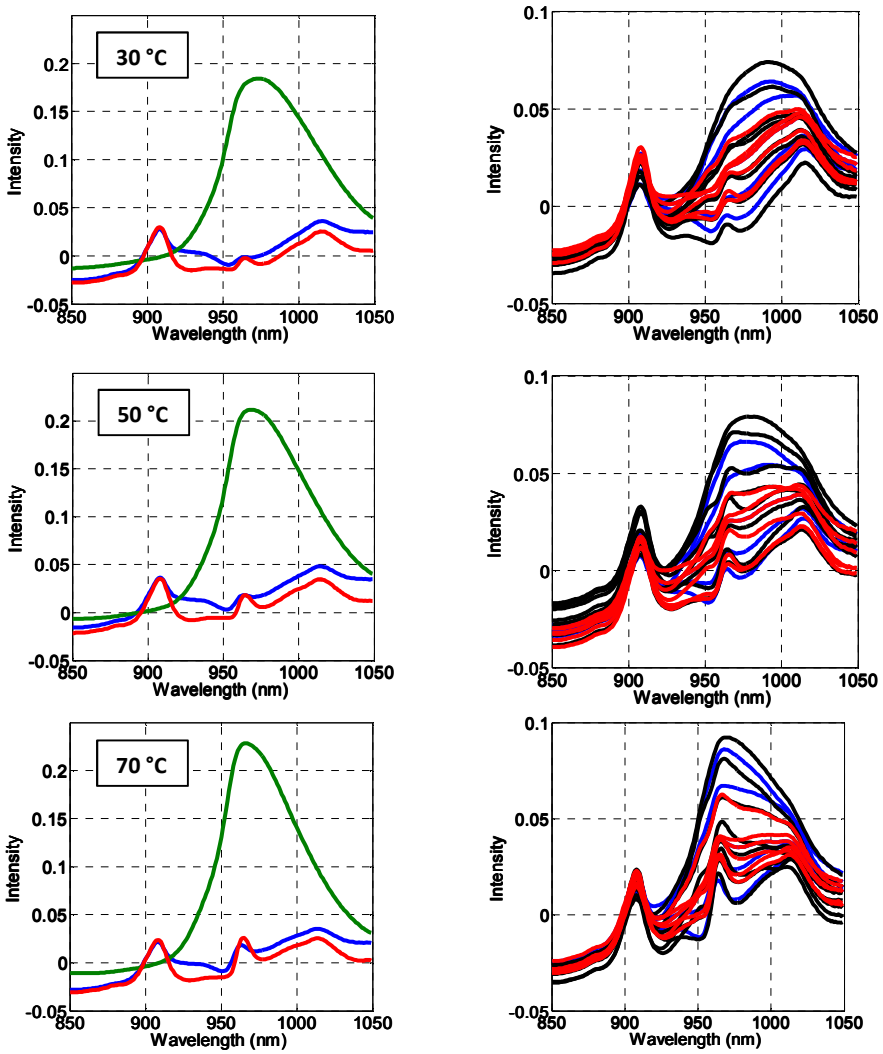

Fig. S-2. Pure component spectra on the right (blue = ethanol, green = water, and red = isopropanol) and on the right are the RR calibration (blue), validation (red) and non-analyte matrix **N** (black) that were also used in the RR calibration. Temperatures are indicated in plots.

Table S-2
Temperature Concentration Design
(Mole fraction).

Sample type	Ethanol	Water	Isopropanol
Cal	0.6715	0.1631	0.1654
Cal	0.6663	0	0.3337
Cal	0.4998	0.5002	0
Cal	0.5003	0	0.4997
Cal	0.1663	0.6669	0.1668
Cal, Non-analyte[a]	0.3332	0.6668	0
Cal, Non-analyte[a]	0.3324	0.5003	0.1672
Cal, Non-analyte[a]	0.3328	0.3340	0.3331
Cal, Non-analyte[a]	0.3222	0.1655	0.5123
Cal, Non-analyte[a]	0.3351	0	0.6649
Cal, Non-analyte[b]	0	0.6671	0.3329
Cal, Non-analyte[b]	0	0.4997	0.5003
Cal, Non-analyte[b]	0	0.3339	0.6661
PC Ethanol	1	0	0
PC Water	0	1	0
PC Isopropanol	0	0	1
Val	0.6644	0.3356	0
Val	0.5003	0.3330	0.1667
Val	0.4994	0.1672	0.3334
Val	0.1670	0.5000	0.3330
Val	0.1662	0.3331	0.5006
Val	0.1622	0.1630	0.6748

[a]Non-analyte samples as approximate constant analyte and hence, mean centered before being placed in **N**.

[b]Non-analyte samples as blanks.