

Near infrared spectroscopy to explore water structure modification induced by filtration processes

Marina Cocchi^{1*}, Mario Li Vigni¹, Stefania Vero², Tiziana M.P. Cattaneo^{2,3} and Vittorio Elia⁴

¹Dept of Chemistry, University of Modena and Reggio Emilia, Modena, 41125, Italy

²CRA-FLC, Research Centre for Fodder Crops and Dairy production, Lodi, 26900, Italy

³CRA-IAA, Research Unit for Agro-Food Processes, Milan, 20133, Italy

⁴Department of Chemistry, University of Naples "Federico II", Naples, 80126, Italy

*Corresponding author: marina.cocchi@unimore.it

Introduction

Studies of very dilute solutions (i.e. containing solutes in concentration below 10^{-7} M) and pure water subjected to physical treatments have recently highlighted that the structure of water is still an open research field. It is well known that the formation of hydrogen bonds is primarily responsible for the peculiar properties of water, such as its high boiling point or its density variation with temperature. However, unexpected results have been observed on extremely dilute solutions perturbed by mechanical treatments and analysed with calorimetric techniques, as well as the formation of nanoclusters obtained by evaporation at room temperature and atmospheric pressure of solutions of NaCl in concentration 10^{-7} M which highlight that not all that concerns water is yet clear.¹⁻³

It has been observed that water conductivity varies after filtration or sterilisation by more than is accounted for by natural atmospheric contamination of distilled water alone. This discrepancy has recently been ascribed to the formation of aqueous nanostructures.⁴ Elia et al.⁵ has suggested that the formation of dissipative structures accounts for the difference observed. In recent years, near infrared (NIR) spectroscopy has been applied to monitor changes in water structure induced by solutes and/or temperature variation - this field of research has been called aquaphotomics. This technique seems to be the most appropriate approach for studying the formation of nanostructures in water. The present study was focused on the variations induced by filtration processes on pure water analysed by NIR and chemometrics.

Materials and Methods

Samples

The sample set consisted of 61 samples: 10 samples of deionised water that were not filtered (reference water), 38 samples divided into two subsets of 17 (set A1) and 21 (set A2) that were filtered with the same protocol but obtained in two different experiments, and 13 samples processed with a different protocol (set B).

Near infrared spectroscopy

NIR spectra were acquired at a controlled temperature of 40°C by a NIRFlex N500 (BÜCHI Italia, Assago, Milano, Italy) in the 4000–10000 cm^{-1} range. Spectra were the average of 32 scans and were collected at 8 cm^{-1} resolution and an optical path length of 1 mm. Only the 5400–10000 cm^{-1} region was considered in data analysis due to the high absorbance of the solution in the 4000–5400 cm^{-1} region.

Data pre-treatment

Spectra were mean centred before data analysis. In addition, the baseline was removed by using a weighted least squares (WLS) baseline algorithm; this uses an automatic approach to determine which points are most likely due to baseline alone and avoids the creation of highly negative peaks.

Data analysis

Principal component analysis (PCA) was used as an exploratory data analysis tool to see how the filtered water solutions clustered with respect to the reference water samples. To see if discrimination could be achieved, the WILMA-GA algorithm⁶ was employed. This algorithm operates feature selection in the wavelet domain by using a genetic algorithm⁷ for selecting the most discriminant wavelet coefficients according to the minimum root mean square error in cross-validation criterion obtained by partial least squares discriminant analysis (PLS-DA). According to the algorithm each signal is decomposed, by applying the wavelet transform, to the maximum allowed decomposition level and all the coefficients obtained by this decomposition are considered. However, in order to avoid over-fitting the GA selection procedure is

Reference paper as:

M. Cocchi, M.L. Vigni, S. Vero, T.M.P. Cattaneo and V. Elia (2012). Near infrared spectroscopy to explore water structure modification induced by filtration processes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGovern, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 243-246.

completed in two steps: firstly the coefficients are grouped in intervals and the maximum number of intervals is kept below 250; then up to 250 of the most frequently selected coefficients are considered for the second selection step.

The selected coefficients are used to build the final PLS-DA model and may be reconstructed in the original signal domain, by using the inverse wavelet transform, to highlight the selected spectral regions. Different wavelet filters were tested and the best results were obtained with a db2 wavelet filter.

In order to validate the PLS-DA model, three reference water samples and seventeen filtered water samples taken from set A1, A2 and B, have been randomly assigned to the test set. Thus, the training set consisted of 7, 11, 14 and 9, reference, set A1, set A2 and set B samples respectively.

Results

Results from PCA of all data sets showed that filtered water samples (sets A1, A2 and B) were well separated from reference water samples (scores plot of the first two PCs, Figure 1). The two categories are discriminated mainly on PC1 and the corresponding PC1 loadings (Figure 2) show a maximum at about 6948 cm^{-1} ; this absorption will be more intense for filtered water, showing positive PC1 scores values. PC2 loadings (Figure 2) show a behaviour which is typical when horizontal shift occurs. In this case, it has been interpreted as an indication of the presence of a convolved band at about 6892 cm^{-1} that changes position between the two categories, the interval being $6772\text{--}7072\text{ cm}^{-1}$. This may be attributed to different hydrogen bond networks.

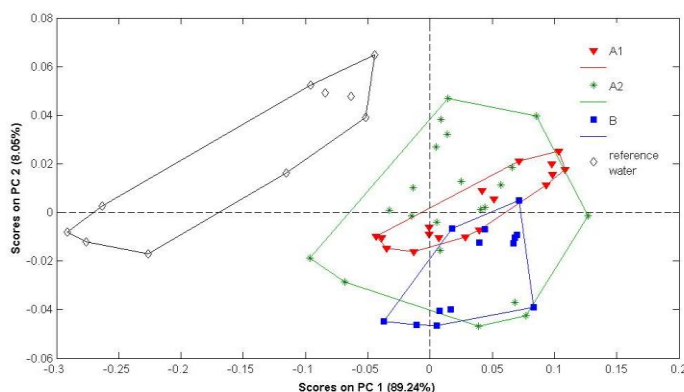


Figure 1. PC1 vs PC2 scores plot from PCA of all samples. The different sets are highlighted by polygons: open diamonds - reference water; red triangles - set A1; green asterisks - set A2 and blue squares - set B.

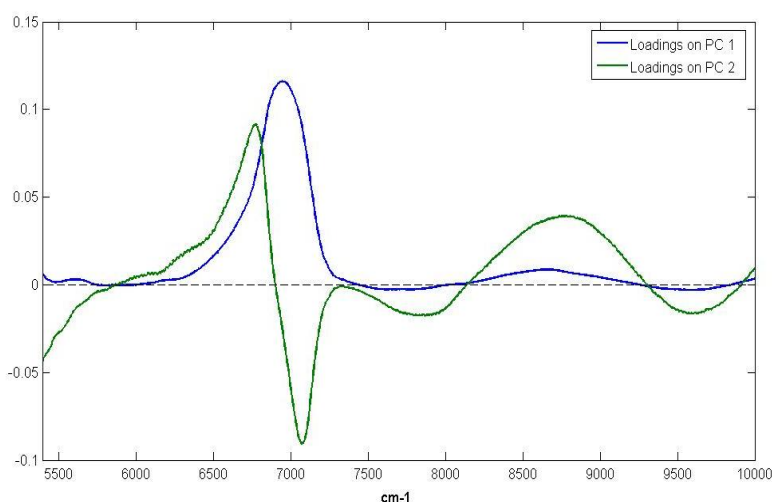


Figure 2. Plot of PC1 and PC2 loadings vs. wavenumber (cm^{-1}) from PCA of the total dataset.

The WILMA-GA algorithm was applied and the PLS-DA model based on selected wavelet coefficients showed no classification error for either both category (Figure 3). To assess the relevance of each wavelet

Reference paper as:

M. Cocchi, M.L. Vigni, S. Vero, T.M.P. Cattaneo and V. Elia (2012). Near infrared spectroscopy to explore water structure modification induced by filtration processes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGovern, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 243-246.

coefficient for discrimination, the values of the variable importance in projection (VIP) parameter and the regression coefficient values were used. Then, the resulting most significant coefficients for discrimination have been reconstructed in the original domain, i.e. by using the inverse FWT it is possible to come back to the wavelength domain and see which is the “spectral profile” depicted by a given wavelet coefficient or a set of them. The wavelet coefficients bearing positive signs in the PLS-DA model, hence directly correlated to reference MilliQ water samples (we set the y values equal to one for this class) were reconstructed separately from those having a negative sign (hence directly correlated with filtered water). The reconstructed signals are shown in Figure 3b and Figure 3c, respectively. Since the selected wavelet coefficients belonged to details (high frequency content of the analysed signals) they look rather different from the original signals and thus the raw signals (scaled by a factor of ten for graphical representation in the same plot) have been also reported in the same figures.

Considering the relevant spectral ranges for filtered water (Figure 3c) 6900-6940 cm^{-1} and 7096-7128 cm^{-1} , it is confirmed that there is a variation in the hydrogen bond networks and thus water clusters form, induced by the filtration process.

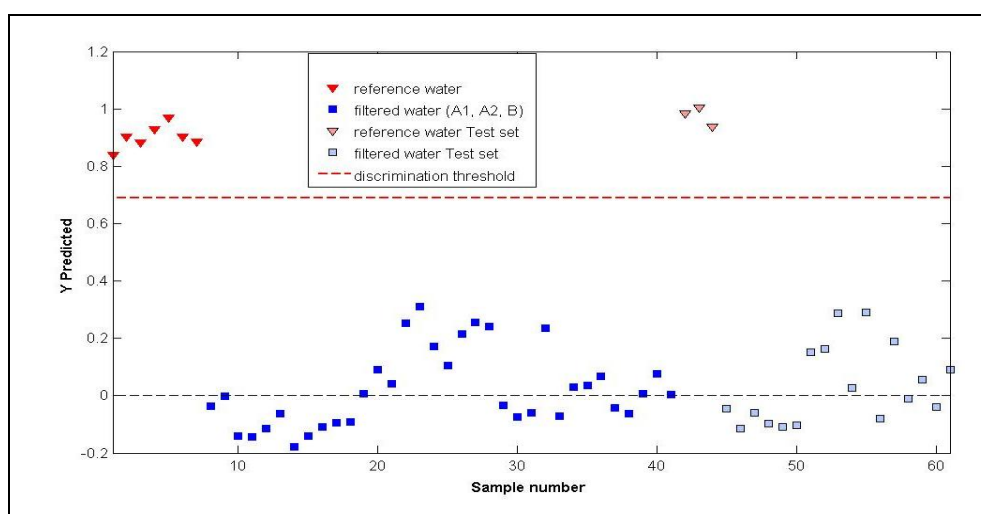


Figure 3a. Y predicted vs. sample number. PLS-DA model calculated on selected wavelet coefficients. Reference water samples were assigned a nominal Y value of 1 and filtered water samples 0. The dashed red line represent the class1 (reference water) assignment threshold calculated in cross-validation.

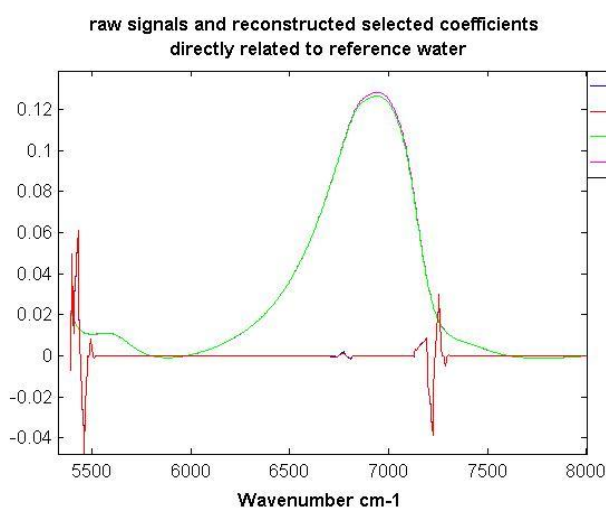


Figure 3b. Raw signals and signals reconstructed by using only selected wavelet coefficients related to reference water samples, i.e. showing positive values of PLS-DA regression coefficients

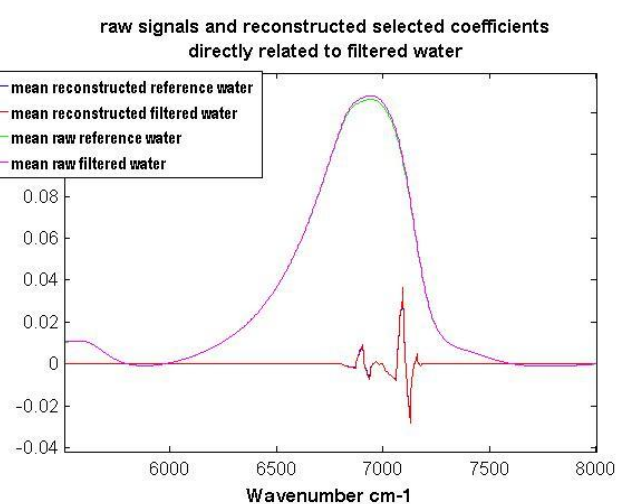


Figure 3c. Raw signals and signals reconstructed by using only selected wavelet coefficients related to filtered water class, i.e. showing negative values of PLS-DA regression coefficients.

Conclusion

New insight into water structure variation induced by filtration processes has been gained by near infrared spectroscopy coupled to chemometrics. Feature selection in the wavelet domain assisted data interpretation. In particular, the filtration process seemed to alter the hydrogen bonding pattern of water molecule clusters. Chemometrics pointed to discriminant spectral bands and allowed rationalisation in terms of hydrogen bond pattern variation.

Acknowledgements

A particular acknowledgment to SISNIR (the Italian Society for Near InfraRed Spectroscopy) for travel grants which enabled the participation of Marina Cocchi and Mario Li Vigni in NIR2011.

References

1. V. Elia, L. Elia, M. Marchese, M. Montanino, E. Napoli, M. Niccoli, L. Nonatelli and F. Savarese, *J. Mol. Liq.* **130**, 15-20 (2007).
2. S.Y. Lo, X. Geng and David Gann, *Phys. Lett. A* **373**, 3872-3876 (2009).
3. R. Giangiacomo, *Food Chem.* **96**, 371-379 (2006).
4. L. Montagnier, J. Aïssa, S. Ferris, J. L. Montagnier and C. Lavallée, *Interdiscip. Sci. Comput. Life Sci.* **1**, 81-90 (2009).
5. V. Elia, E. Napoli and M. Niccoli, *J. Mol. Liq.* **149**, 45-50 (2009).
6. M. Cocchi, C. Durante, G. Foca, M. Li Vigni, R. Leardi and A. Ulrici, In VI Colloquium Chemometricum Mediterraneum, St. Maximim La Sainte-Baume, France, 5-7 September 2007, p. 28 (2007).
7. R. Leardi, *J. Chemom.* **14**, 643-655 (2009).