TONY DAVIES COLUMN

Svante Wold 1941–2022

Antony N. Davies

SERC, Sustainable Environment Research Centre, Faculty of Computing, Engineering and Science, University of South Wales, UK

It is a very rare event when someone comes along and changes the everyday language we use. This short article is to mark the passing of Svante Wold on 4 January 2022 who holds that honour!

Chemometrics

The word Chemometrics is a lovely word merging Chemistry and Statistics—perfectly describing at a high level what has essentially become a scientific discipline in its own right. One that has spawned many great careers and advances in the way we actually do chemistry and spectroscopy.[1]

Family heritage

Svante Wold wasn't the first statistically talented person in the Wold family, his father Herman Ole Andreas Wold was a famous statistician in his own right, born in Skien, Norway on Christmas Day 1908, Herman and his family emigrated to Sweden where they settled. Herman Wold may well be best remembered as a pioneer of Partial Least Squares (PLS) modelling, although he worked in the field of economics and the analysis of data where short-term fluctuations may hide key longer-term changes in the data.[2] Svante's mother was the mathematician Anna-Lisa

Arrhenius Wold, the daughter of Svante Arrhenius who, although a physicist by training, became world renowned for his research in the field of physical chemistry, was the winner of the Nobel Prize for Chemistry in 1903 and the first person to use scientific principles of physical chemistry to look at the relationship between atmospheric CO_2 and global warming.[3]

I have no idea if this family history influenced Svante Wold's decision making around his own career path, but certainly having parents and grandparents who were radical ground breakers in their own fields must have played a role. We can only be very thankful that Svante decided to work in the field of chemical data processing and provided us with the term "Chemometrics". My previous column co-editor, A.M.C. Davies, remembers Svante as a very friendly personality and was fortunate enough to dine with him during a visit.

SIMCA (the statistical model)

Classical hard modelling of data looks to separate data into specific classes for analysis and prediction of, for example, properties based on new data falling into one class or another. A very simplistic binary approach and often hard to apply to data from far more complex real-world systems. Svante proposed a more "soft modelling" approach to data analysis which better captures what we see in chemistry and spectroscopic analyses, and called it SIMCA for Soft Independent Modelling of Class Analogies. This allows data to be statistically analysed for classification and placed into one class, or two, or none. Figure 1 provides a primitive representation showing classes overlapping.

Richard Brereton uses the analogy of when a spectroscopic measurement on a series of chemical entities clearly identifies
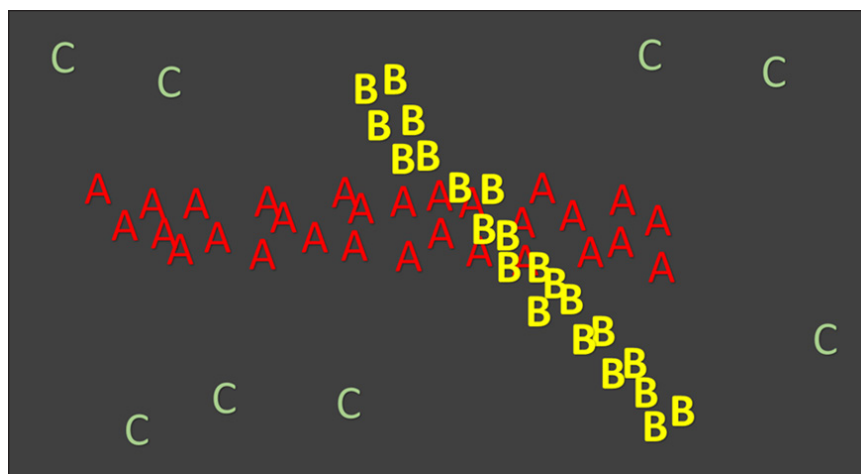
Figure 1. SIMCA analysis is a "soft" modelling technique allowing classes to overlap.
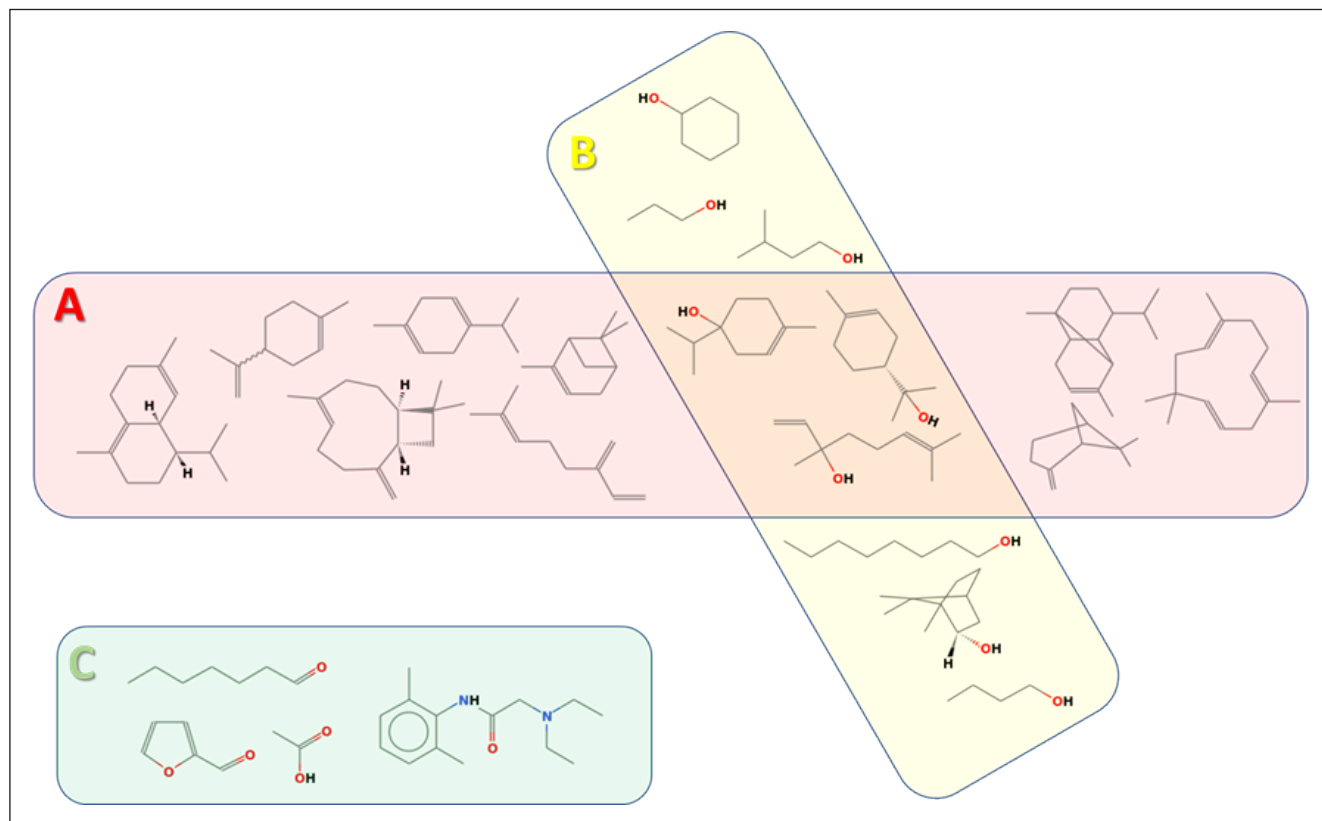
Figure 2. Classification of samples based on spectroscopic evidence of functional group chemistries.

the presence of alkenes as well as esters. Clearly spectroscopic analyses could place molecules in one class or the other, both or neither reflecting on the functional group distribution in the measured sample molecules. Figure 2 shows such a soft modelling classification example on the spectroscopic analysis of various analytical samples containing alcohols, alkenes and other analytes.

In our own work, the power of this soft modelling approach to sample classification is regularly used and was demonstrated when we looked at deliberate adulteration of olive oils with sunflower oil. Principal component analysis of the Raman spectroscopic data, amongst others, clearly separated samples of pure Peloponnese olive oils and samples adulterated with only 2 % sunflower oil, Figure 3. A Coomans plot is a simple way to display classification results by

dividing the plot into four quadrants. The top-right quadrant is where samples which have not been classified into one of the two classes plotted are shown, bottom-left show samples which have

been assigned membership of both classes, top-left and bottom-right to one class or the other. For a more detailed discussion see an earlier column by A.M.C. Davies and Tom Fearn.[6]
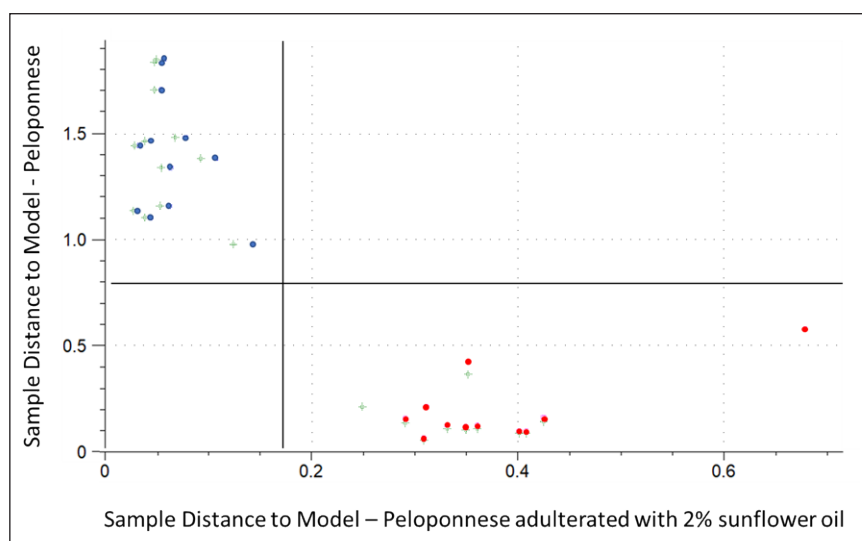

Figure 3. Coomans plot of SIMCA classification of Greek olive oils, pure and adulterated with 2 % sunflower oil (adapted from Reference 5).

SIMCA (the commercial software package)

Svante was also involved in the commercialisation of chemometrics tools with the founding of Umetrics in 1989 with Asa Nilsson, Conny Wikstrom and Rolf Carlsson. Confusingly, the software package SIMCA was developed and marketed which contained more analytical tools than just SIMCA (the modelling approach). The company was successful and in 2017 was purchased by Sartorius with whom they had been collaborating for around five years. Sartorius purchased the company for US$72.5 million from the US MKS Instruments Group.

With Svante Wold's passing we have lost a founding father of the field of Chemometrics. Most people know that the Chemometrics name was first used in the 1971 grant application. When they met, A.M.C. Davies had the cheek to ask his dining partner if the original seminal 1971 grant application had actually been successful? Svante happily told him it had! Where would be now had the grant application been turned down?

References

1. S. Wold, "Chemometrics; what do we mean with it and what do we want from it", *Chemometr. Intell. Lab. Syst.* **30,** 109–115 (1995). https://doi.org/10.1016/0169-7439(95)00042-9
2. R.G. Brereton, *Chemometrics, Data Analysis for the Laboratory and Chemical Plant*. John Wiley & Sons (2003). https://doi.org/10.1002/0470863242
3. S. Arrhenius, "On the influence of carbonic acid in the air upon the temperature of the ground", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **41(251),** 237–276 (1896). https://doi.org/ff3hbv
4. S. Wold and M. Sjostrom, "SIMCA: A method for analyzing chemical data in terms of similarity and analogy", in *Chemometrics Theory and Application*, Ed by B.R. Kowalski. American Chemical Society Symposium Series 52, American Chemical Society, pp. 243–282 (1977). https://doi.org/10.1021/bk-1977-0052.ch012
5. A.N. Davies, P. McIntyre and E. Morgan, "A study of the use of molecular spectroscopy for the authentification of extra virgin olive oils: Part 1: Fourier transform Raman spectroscopy", *Appl. Spectrosc.* **54(12),** 1864–1867 (2000). https://doi.org/bn46qz
6. A.M.C. Davies and Tom Fearn, "Back to basics: multivariate qualitative analysis, SIMCA", *Spectrosc. Europe* **20(6),** 15–19 (2008). https://doi.org/10.1255/sew.2008.a1

Tony Davies is a long-standing *Spectroscopy Europe* column editor and recognised thought leader on standardisation and regulatory compliance with a foot in both industrial and academic camps. He spent most of his working life in Germany and the Netherlands, most recently as Lead Scientist, Strategic Research Group – Measurement and Analytical Science at AkzoNobel/Nouryon Chemicals BV in the Netherlands. A strong advocate of the correct use of Open Innovation.

https://orcid.org/0000-0002-3119-4202
antony.n.davies@gmail.com