**Solvent Selection Methods and Tool.**

Patrick M. Piccione1,\*, Julia Baumeister1, Thomas Salvesen2, Yannick Flores1, Christophe Grosjean1, Vikrant Murudi4, Ashok Shyadligeri4, Olga Lobanova3, Christian Lothschütz1,†

*1 Syngenta, 4333 Münchwilen, Switzerland; 2 Syngenta, 1870 Monthey, Switzerland; 3 Syngenta, Bracknell, RG42 6EY, UK; 4 Syngenta, Corlim, Goa, 403110, India † Current address: DSM, 4334 Sisseln, Switzerland*

*\*Corresponding author:* [*patrick.piccione@syngenta.com*](mailto:patrick.piccione@syngenta.com)

**Highlights**

* A solvent selection tool developed at Syngenta is described through examples.
* Its guiding principles are ease of use, flexibility, and empowerment of chemists.
* Its interactive visualizations help exploring the solvent parameter space efficiently.

**1. Introduction**

Chemical manufacturing processes, especially for fine chemicals, often take place in the liquid phase. Solvents facilitate reactions, enable separations, and lower viscosity for convenient transport. Where possible, water is typically used, but many phenomena require different properties, better provided by organic solvents. The multitude of possible functional groups leads to even more numerous property combinations. Solvent selection is thus an important decision in chemical process development. Many methods have been put forward to optimize this selection, from high throughput screening to predictive models. The latter typically rely on physicochemical descriptors and can be made more accessible by embodiment in software tools [1, 2]. To get the most value from such tools, the corresponding theories and assumptions must be understood. To achieve both prediction and insight, an in-house tool was built at Syngenta and is described here.

**2. Methods**

Solvent selection methods have been researched by chemists as well as process systems engineers [3, 4]. Information technology skills are required to produce robust, attractive, and sustainable tools. A multi-functional project team was thus assembled (see Table 1). The basic philosophy was to build a tool empowering trained chemists to explore the process space – not a prescriptive tool with definite decisions, but a *suggestive* one. A main use thereof is defining the next iteration of solvents to try in the laboratory.

|  |  |  |  |
| --- | --- | --- | --- |
| **Chemists: Main user community** | **Process engineers** | **Statisticians** | **Data scientists** |
| Definition of usage criteria  Data for repository | Theories for property/phenomena correlations  Functionalities and workflows | Fitting techniques | IT implementation |

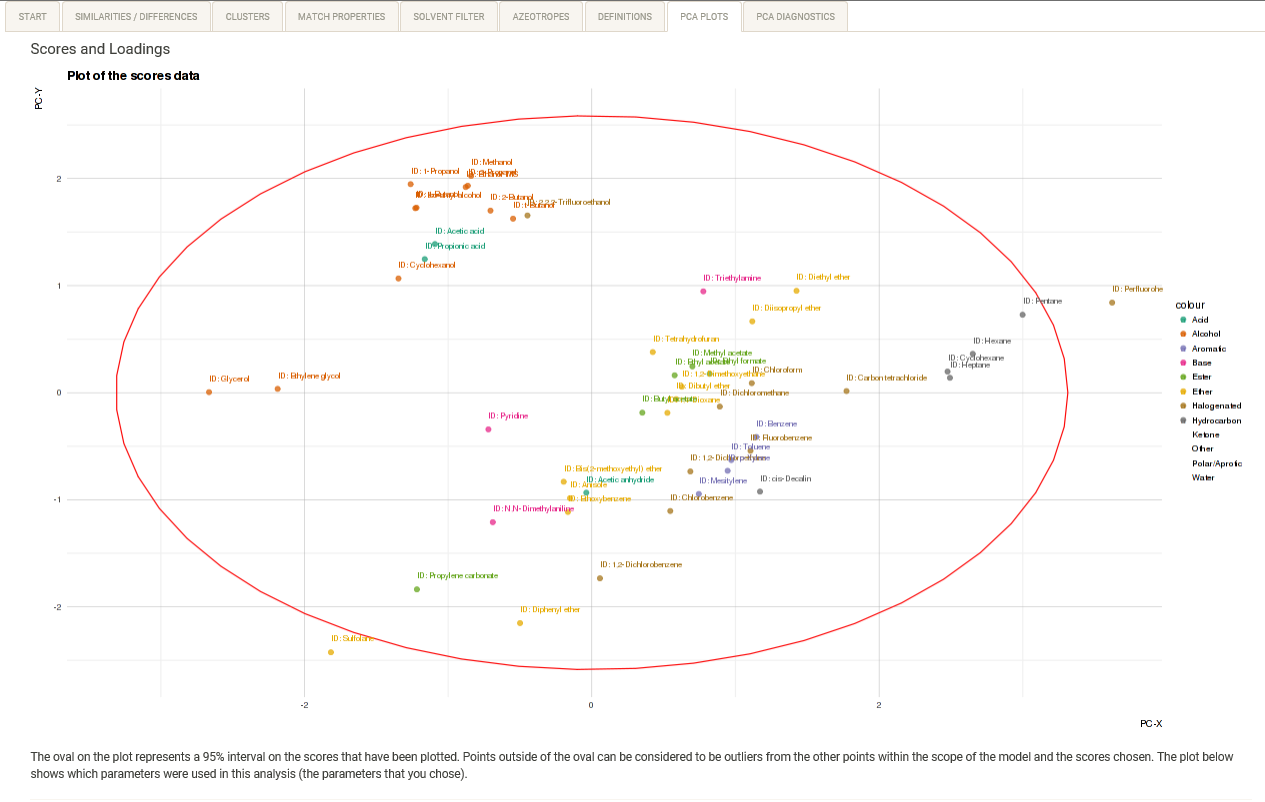
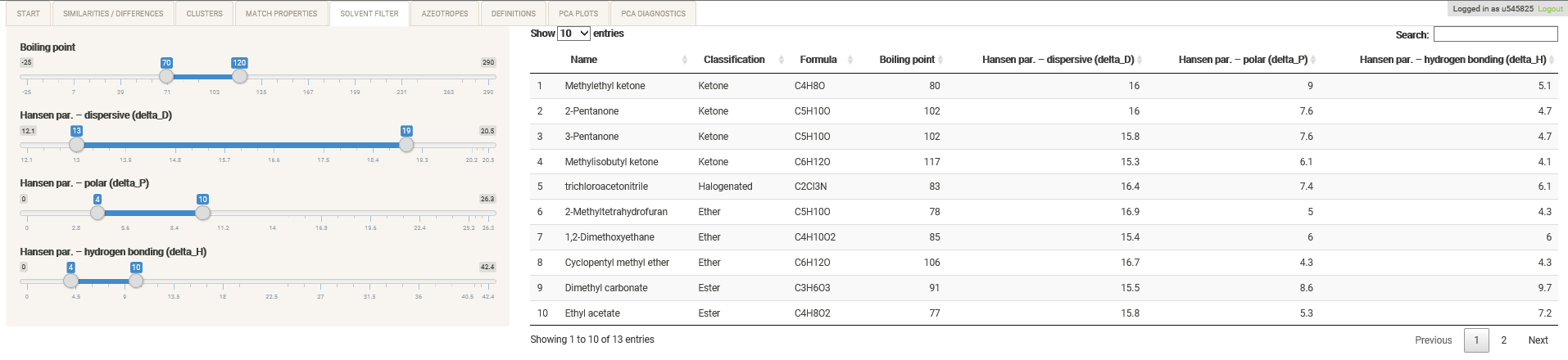
**Table 1.** Experts’ roles on the project team

Physical and other properties were compiled for approximately 130 solvents in a spreadsheet (“the database”), from the following broad categories: identifiers; measured properties; physicochemical derived descriptors; HSE scores. The following calculations were then automated in the software R-Shiny: 1) similarity and difference searching using hierarchical cluster analysis; 2) multi-criteria searching in property value space; 3) principal component analysis (PCA) for visualization as maps. All these functions allow dynamic, real time selection of the properties of interest by the user.

To underpin the *use* of the tool, several theories were used; in particular, solvatochromic / Abraham [5] and Kamlet-Taft [6] descriptions. The corresponding sets of properties were selectable together for convenient access to the occasional user.

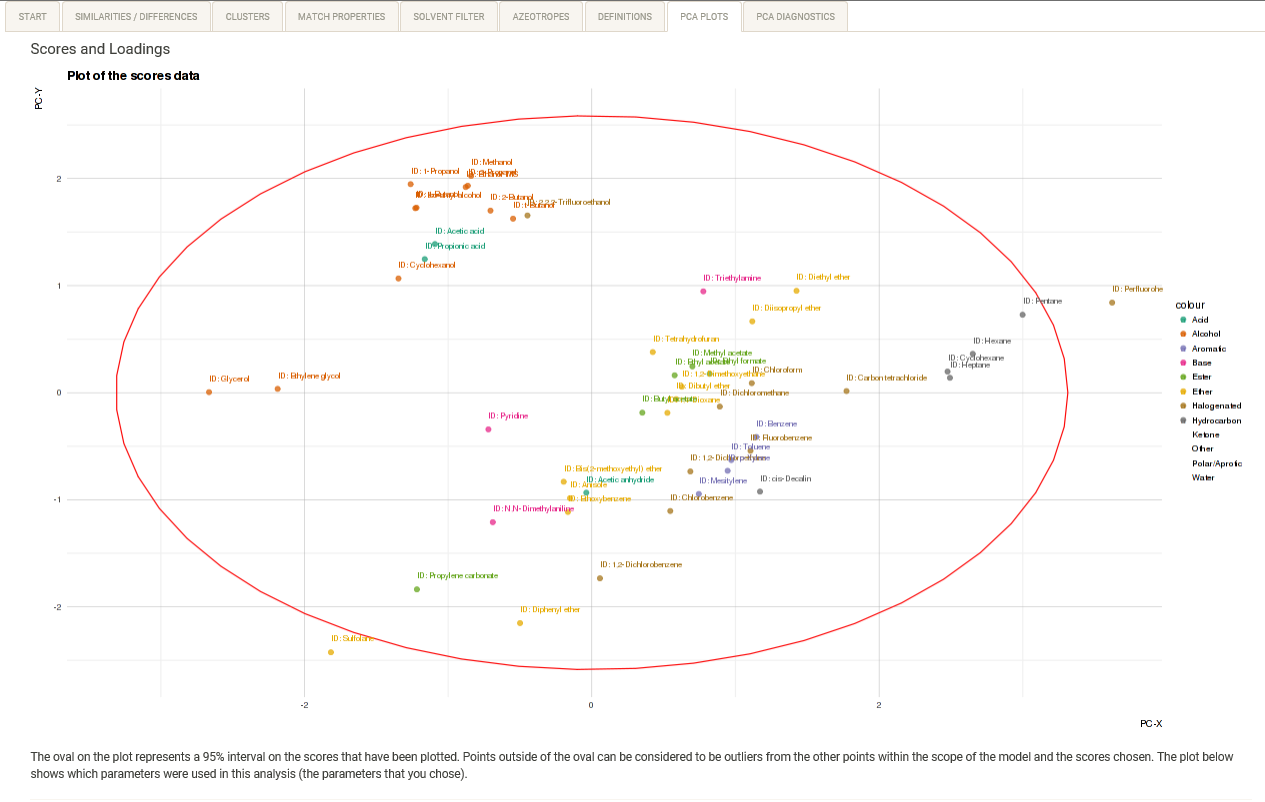
**3. Results and discussion**

The R-Shiny tool appears as a typical browser interface. Parameter and visualization choices are set in a “start” tab, leading to real-time calculations in various results tabs. The tool is designed to be usable in many different ways. Cluster analyses yield solvents similar, to, or different from, a specific one. Multidimensional filtering allows several property ranges (e.g., solubility parameters and boiling point) to be matched simultaneously (Fig. 1a). PCA maps based on parameters known to matter for a process (e.g., Kamlet-Taft α, β, π\* + boiling point, Fig. 1b) also provide alternatives.



**a**

**b**



**Figure 1.** Screenshots of the matching (a) and PCA (b) tabs in the tool.

**4. Conclusions**

Different scientific specialisms lead to different languages and approaches. Tool development was found to be more intrinsically iterative than originally expected. While physical properties are a Rosetta stone for the meeting of chemists, engineers, and statisticians, progress required real usage cases. It was far easier to discuss mock-ups and share user stories than to write requirement specifications, leading to the mantra “show, don’t tell”. To remain meaningful to physical scientists, graphs and points required extensive labelling. Several concrete examples of use are documented in [7]. Possible improvements were continuously identified. Two important directions are: further incorporation of chemical theories; increased tool interactivity and filtering capability.

**References**

[1] L.J. Diorazio, D.R.J. Hose, N.K. Adlington, Org. Proc. Res. Dev., 2016, 20, 760–773.

[2] C.M. Alder, J.D. Hayler, R.K. Henderson, A,M. Redman, L. Shukla, L.E. Shuster, H.F. Sneddon, Green Chem., 2016, 18, 3879-3890.

[3] D. Curzons, D.J.C. Constable, D.N. Mortimer, V.L. Cunningham, Green Chem., 2001, 3, 1–6.

[4] P.M. Murray, F. Bellany, L. Benhamou, D.-K. Bučar, A. Tabor, T.D. Sheppard, Org. Biomol. Chem. 2016, 14, 2373-2384.

[5] M.J. Kamlet, J.-L.M. Abboud, M.H. Abraham, R.W.Taft, J. Org. Chem. 1983, 48, 2877-2887.

[6] M.H. Abraham, Chem. Soc. Rev. 1993, 73-83.

[7] P.M. Piccione, J. Baumeister, T. Salvesen, C. Grosjean, Y. Flores, E. Groelly, V. Murudi, A. Shyadligeri, O. Lobanova, C. Lothschütz, Org. Proc. Res. Dev. 2019, 23, 998-1016.