Brazilian Journal of Analytical Chemistry
2021, Volume 8, Issue 32, pp 6-12
doi: 10.30744/brjac.2179-3425.interview.ajcapdevila

BrJAC

INTERVIEW

Anna de Juan, an internationally recognized researcher on Chemometrics, spoke to BrJAC

Anna de Juan Capdevila
Associate Professor
Faculty of Chemistry at the University of Barcelona, Spain

Professor Anna de Juan's research focuses on the theoretical development of multivariate curve resolution and its application to process analysis, hyperspectral image analysis, and general analytical and bioanalytical problems. She has published more than 140 works (h index = 40) and has given more than 200 presentations at international conferences. She received the 4th Chemometrics Elsevier Award in 2004 and the Kowalski Prize from the *Journal of Chemometrics* for the best applied paper in 2009. She has served on the Editorial Advisory Board of *Chemometrics and Intelligent Laboratory Systems* since 2002 and of *Analytica Chimica Acta* since 2006. Her teaching activity covers undergraduate and graduate topics related to chemometrics and analytical chemistry, and she has been an invited professor for short periods at the Université de Lille (France), the University of Dalhousie (Canada), the Università di Modena e Reggio Emilia (Italy), the Institute of Advanced Studies in Basic Sciences (IABS) Zanjan (Iran), the Universidad de Santa Fe (Argentina), and the Universidad Pontificia de Valparaíso (Chile).

The first question I would like to ask you is: Why science? There are plenty of opportunities and possibilities to make a life. So, what were the factors that lead you to chemistry, what fascinated you?

To be honest, when I was a kid, I was fascinated by many areas of knowledge and arts. I loved science, but also enjoyed learning languages and had a strong artistic inclination that made me do classical dance for many years, for instance. I think I could have been very happy dedicating my life to do language research or to an artistic activity. In secondary school, I discovered chemistry and, to some extent, this is the most "magic" science among all others; it connected a bit with the artistic/fantasy side that I have always had. I met a very special chemistry teacher in high school, doña María, not the best in terms of conceptual explanation, but her passion when talking about chemistry captivated me. Chemistry experiments were also very exciting, because you could "see" in a spectacular way sometimes, the phenomena you were told about. At university, I realized that it was a good choice and I could discover the many real sides of chemistry that always connect with the understanding of any event in nature and life, providing explanations at all possible scales, from the molecular theory to the macroscopic observation.

Cite: Capdevila, A. J. Anna de Juan, an internationally recognized researcher on Chemometrics, spoke to BrJAC. *Braz. J. Anal. Chem.,* 2021, *8* (32), pp 6-12. doi: http://dx.doi.org/10.30744/brjac.2179-3425.interview.ajcapdevila

How was the beginning of your career? How did you discover chemometrics and why were you interested in this field? Could you please comment about the time you were at Vrije Universiteit Brussel (Belgium), in Luc Massart's laboratory? At the same time, Prof. Ronei Jesus Poppi was there as a post-doc student. Did you get to know each other as colleagues or even worked together?

When I started my MSc, I realized I could not see myself doing routine work for the rest of my life and decided to start a research career. Professor Enric Casassas influenced significantly this choice. He was a great researcher and a multitalented person, with a strong Renaissance wisdom spirit. With him, I learned how to approach a scientific problem and the importance of conveying appropriately scientific knowledge, writing and presenting science using a precise language and clear and attractive images and explanations. He introduced me to Romà Tauler, a young professor who had been supervised by him, and suggested that I could do "something" with him. The "something" was chemometrics. So, in the early 90s, at Romà's initiative, I started reading the book *Factor Analysis in Chemistry* by Malinowski and to do some basic works with him using factor analysis and target factor analysis to study solvent effects in equilibria. At that time, in 1992, I attended a chemometrics school in Ulvik (Norway), organized by Kvalheim's team, and I attended a COMETT school in Santiago de Compostela in 1993, a great initiative supported by the European Union, where the best chemometric experts in Europe were teaching and offering a wide vision of chemometrics. There, Bernard Vandeginste, a teacher in the two schools mentioned, played a crucial role for my later stage in Massart's laboratory. I arrived at Massart's laboratory in 1995, with fundamentals of factor analysis and a good knowledge of the developments in multivariate curve resolution. That was a very exciting period in my life and served to consolidate the choice of chemometrics as my research field. Massart's lab at the time comprised around 20 people between PhD students, post-doctoral people, and invited professors (the essential role of people like Beata Walczak through many years has to be mentioned) from all over the world. Continuous group meetings were taking place and almost all possible areas within the chemometric field were addressed by someone in the group. Theoretical developments were accompanied by practical, challenging problems brought to the laboratory by the industrial consortium (ChemoAC) associated with the group. My personal contribution was on the study of a pure variable selection method, the needle algorithm, on the design of some new constraints for multivariate curve resolution and on the comparison of multiset analysis with other multi-way algorithms, together with Sarah Rutan, an invited professor at that time. As a scientific and multicultural personal experience, it was a unique year in my life. During my period there, I met Ronei Poppi, who joined the group for a research stage. Paula Fernandes de Aguiar, then a PhD student in Massart's lab, introduced him to us. Although I could not work with him directly, he was the kind of person appreciated by everyone: shy, but extremely gentle and clearly a very hard worker, trying to learn as much as he could to bring it back to Brazil and create some years later not only a group, but a continuous inspiration for the development of chemometrics in his country.

You have given a brilliant contribution in the field of multivariate curve resolution for different purposes and combined with different analytical techniques. What are your lines of research and main interests at the moment? What has changed in your interests since the beginning of your career?

Right now, my main interest is still mainly in multivariate curve resolution both from a theoretical and practical side. In terms of theoretical aspects, I am very much interested in the adaptation of the algorithm to new data structures, modifying the algorithm or the underlying mathematical models used. This implies pursuing the use of incomplete multisets trying to understand better how the different factorizations can be addressed and combined and how these structures and data analysis compare with other approaches developed with the same objective, such as combined tensor and matrix factorizations. Another side is insisting on using the value of the scores and loadings obtained from multivariate curve resolution as seeding information for other data analysis purposes, e.g., heterogeneity studies, segmentation, clinical/food sample profiling, etc. We all use score/loadings from principal component analysis in a very natural

way as compressed expressions of the information in our data, but do not think that meaningful multivariate curve resolution scores and loadings can be equally useful and provide an easier interpretation of the results obtained.

The main applied field now is the application and adaptation of multivariate curve resolution to hyperspectral image analysis. This field has new challenges, such as the active treatment of the spatial information provided by the image, and the data fusion challenges that force us to envision new strategies to cope with the multiscale spatial differences among platforms and the specificities in terms of dimensionality and the underlying model of the spectroscopies that go with them. The massive size of these data is another element, which is common to other kinds of data sets, and that needs to be addressed. The industrial process application is something that we have also tackled in recent years, where sensor fusion for process understanding and process control is a need and a practical introduction of image-based monitoring and control can provide new and more useful results.

For you, what have been the most important achievements in the analytical chemistry recently? Could you briefly comment on recent advances and challenges and how chemometrics could be helpful?

Analytical chemistry is in constant evolution. On one side, the analytical platforms are evolving in incredible ways and provide possibilities to have an integral description of samples from a qualitative and quantitative chemical/molecular point of view, e.g., the latest high resolution mass spectrometry developments, from a multiresolution spatial point of view, e.g., hyperspectral platforms, and from a temporal dimension, all with ultrafast spectroscopies and methods. The role of nanotechnology in analytical chemistry would also deserve a chapter in itself. At the other end, cheap and fast analytical devices allow in situ or in vivo monitoring of processes, samples, and living systems. The problems addressed, e.g., -omics challenges, environmental monitoring, and bioanalytical studies, have gained in complexity and require fusion of very different kinds of information. Chemometrics plays an essential role to combine and interpret reliably all the information provided. Another perhaps philosophical gain has been understanding that we can talk about analytical science, since the multidisciplinary value of analytical chemistry in many neighboring scientific fields is undoubtable nowadays.

Working at the synchrotron ALBA in 2019 with the research group: Sara Piqueras, Adrián Gómez, and Rodrigo Rocha.

Do you believe that the current graduate programs produce quality researchers in the field of analytical chemistry and chemometrics? In your university, is chemometrics taught at the undergraduate and graduate level?

"...in recent times, my impression is that there is a stronger concern in terms of scrutinizing better which concepts need to be conveyed to prepare both researchers and professionals in graduate programs, and a lot of attention is paid to develop other skills, such as the communication of scientific knowledge and development of critical thinking."

To some extent, I think that education programs have always provided quality researchers; otherwise, we would not be where we are now. But in recent times, my impression is that there is a stronger concern in terms of scrutinizing better which concepts need to be conveyed to prepare both researchers and professionals in graduate programs, and a lot of attention is paid to develop other skills, such as the communication of scientific knowledge and development of critical thinking.

In the Faculty of Chemistry at the Universitat de Barcelona, chemometrics is present in both undergraduate and graduate curricula. In the undergraduate program, a compulsory Advanced Analytical Chemistry topic includes classical statistics, design of experiments, and univariate calibration. In the last year of the Chemistry degree, an optional chemometrics topic offers the fundamentals of multivariate analysis, including theoretical and practical work about exploratory, calibration, and classification methods. Team work on case studies based on real data sets is one of the assets of the subject, where a real problem is posed and chemometric knowledge, team work, and decision-making skills are built. In the graduate program linked to the Master of Analytical Chemistry, a subject called Chemometrics and Process Control is offered. The same contents as in the undergraduate option, incorporating multivariate curve resolution and multivariate statistical process control, are offered.

In the lab: Víctor Olmos, Lorenzo Strani, Sanae Benabou, Silvia Concolino, Adrián Gómez, Alba Navarro, and Rodrigo Rocha.

Do you work in collaboration with industry (chemistry, pharmaceutical, etc.)? How do you see this relationship (both for university funding and for the spread of academic knowledge)?

We have been working more often in contact with industries than in the past. We had a very good opportunity under the frame of the European Union project ProPAT (pro-pat.eu), focused on the development of cheap sensors and related process control methodologies for industrial end-users. To be honest, I think that industries and universities still need to learn from each other. We need to understand and try to adjust to the real objectives of the industrial problems, whereas industries need to understand that perhaps a more open view of the problem can help them to find a more complete solution and that knowledge is value and appropriate funding is required. In the last decades, small technological companies have emerged that act as a bridge between the industrial end-user and the academic world. I think they often play a very interesting role because they have a language understandable by both sides and they can be useful to foster collaboration.

Thinking of chemometrics, the value of the discipline is getting recognized, but I have also realized, through attending different PAT-related meetings, that many industries still tend to see engineers as their

natural partners and chemists or chemometricians as more academic researchers. In this sense, there still exists different paradigms in these two worlds – the deterministic way to interpret processes by engineers and the soft experimental monitoring and control linked to chemometrics – which is sometimes seen as not completely reliable. Proving the usefulness and compatibility of both approaches is needed.

You are a woman in science, we know that historically this hasn't been always easy. Along your scientific career, have you faced any gender prejudice?

"I consistently see more young women in science, and there is a very positive movement all over the world to promote the presence of women in science, as a life option as natural for them as any other one."

Actually, I have never suffered such a problem in my home institution. I grew up in a research group where gender was not an issue, and this surely helped me to feel more confident once I moved to other contexts and situations. Generally speaking, I am really happy with the treatment I have received from my colleagues, although I may refer to very few occasions where I felt uncomfortable as a woman and certain situations where I really had to be more assertive than needed to make my point. But it is true that in many scientific conferences, and chemometrics conferences are not an exception, you can still see not many women giving keynotes or plenary lectures, and this is a problem that comes from many years ago. For too long, science has been seen as a masculine territory, and we still drag this historical misconception, joined to the fact of the hidden brilliant women scientists in the past who are just now starting to be rediscovered and granted the credit they have always deserved. In 2020, the Royal Society of Chemistry made a statement on inclusion and diversity in Chemical Sciences (https://www.rsc.org/news-events/articles/2020/jun/id-joint-societies-statement/). These statements mean that much work still needs to be done. However, I am confident, because I consistently see more young women in science, and there is a very positive movement all over the world to promote the presence of women in science, as a life option as natural for them as any other one.

What is MCR GUI? What is your opinion about software interfaces in chemometrics?

MCR GUI is a free, downloadable graphical interface that tries to incorporate all the solid knowledge on the multivariate curve resolution-alternating least squares method that has been developed through the years with the joint effort of many PhD students and researchers, led by Romà Tauler, and with the contributions by Joaquim Jaumot and myself (for the last version of it, please read: J. Jaumot, A. de Juan and R. Tauler. Chemom. Intell. Lab. Sys., 140 (2015) 1 and go to www.mcrals.info for the download). Much credit needs to be given to Joaquim Jaumot, who has been implementing the interface. In our case, the appearance of the interface has promoted a massive increase in the use of the algorithm. It was sharing the knowledge with non-experienced users, providing an additional method to their chemometric toolbox.

I am really in favor of chemometrics software interfaces. Some experts claim that people misuse them and they are the cause of many sorts of mistakes. In my opinion, good interfaces are the best way to encourage people to use chemometrics in their professional or research activities and an invaluable help for teaching purposes. Of course, the use of any tool requires previous preparation and it is the responsibility of the user to follow the necessary education to be able to apply the interface possibilities properly. The only thing I do not like in certain interfaces is the presence of default options. As I mention often to my students, we never work with "default data sets"; they all have their specific characteristics. Interfaces need to offer different options to work and, at the most, suggestions of use, but leave the final decision-making task to the user.

Recently, the coronavirus pandemic has shaken the world in many ways. How do you think this would affect international scientific conferences—are we moving toward the online conferences era? In 2015, you participated in the II Winter School of Chemometrics, which took place at the University of Campinas (UNICAMP), and Prof. Ronei was one of the organizers. What are your memories of that meeting?

Dinner with the speakers and organizers of the II Winter School of Chemometrics (UNICAMP, 2015).

It is clear that coronavirus has affected many aspects in our lives. It has been like a slap in our face reminding us that we are vulnerable and any abuse we do toward our planet and all the nature within will bounce back and bring this pandemic, or there may be others in the future. I do not think all the conference world will move to an online format in the future. Of course, the situation has reinforced these modalities and they are a temporal solution to the problem. I would make a distinction between huge meetings with thousands of people, where the online format or hybrid models are probably very valid options in the future, and the small conferences, very focused on particular fields, e.g., chemometrics, where the direct contact and personal collaboration are invaluable aspects that cannot disappear. This also applies to teaching: We teach online at the university until we can go back again and, if anything, we can envision some kind of hybrid model in the future, but never fully online.

For instance, I cannot imagine losing the pleasure and atmosphere of the live winter school of chemometrics. In the edition where I participated, I was invited with Doug Rutledge, José Manuel Amigo, and Héctor Goicoechea and, as teachers, we still remember the good time we spent there with Márcia Breikreitz, Ronei Poppi, and all other organizers. I have very good memories of that school because of the wonderful organization and, above all, for the huge number of motivated students from all over Brazil and other places in South America. It is one of the few times I have seen so many young people so eager to learn chemometrics. This is the reward of the good education task made by Ronei Poppi and many other teachers in Brazil with respect to the chemometrics discipline.

Participants of the II Winter School of Chemometrics (UNICAMP, Campinas, 2015).

How do you see the connection between these "new" tools and terms that we see in the media nowadays ("big data," "machine learning," "deep learning," "artificial intelligence," etc.) and chemometrics? Are they an extension of currently used chemometrics methods applied to a higher amount of data, or are they part of a separate science?

All eras need their terms, and the evolution of the massive numerical and non-numerical information nowadays has required new labels. It is also true that the concern about the proper interpretation of information exceeds the chemometrics field and people from other domains look for new names—and that is OK. Often, though, when you dig under these labels, you discover good old friends, like machine learning-labeled works that use principal component analysis, and you have to smile. But it is also true that many other old -metrics, e.g., econometrics, psychometrics, have used principal component analysis as well. So, the method does not belong to anyone. There are many tools in data sciences, and the further they go in terms of applications and domains of knowledge the better. We can always claim that we knew many data analysis tools before these new terms were coined, but we can also learn many more in the future, and this is exciting.

In your opinion, what are the skills that an analytical chemist researcher should have that would be vital to the future? What kind of advice would you give to a newcomer to analytical chemistry and chemometrics?

As I mentioned before, the analytical way of thinking will be essential to address many of the current scientific problems in a multidisciplinary way. Looking at the problem as a whole and considering carefully the nature of the question, of the samples involved, have a solid knowledge about the analytical platforms and the possibilities they can offer, have a complete toolbox of data analysis methods to extract reliable conclusions, etc., are all good skills for the future. Analytical chemists are at the interface of many disciplines, and this is a good value.

I could say a similar thing with respect to chemometrics; a chemometrician does not only have the data analysis tools, but also the chemical knowledge to interpret which ones are appropriate depending on the scientific context and which kind of new developments must be done to adapt to new needs in terms of problem complexity or new measurement characteristics. Again, it is at the interface of mathematics/statistics and applied sciences, becoming an essential piece of the game.

If you were starting over today, what would Dr. Anna de Juan say to the young Anna?

Just go for your goals, it will not always be easy, but it is worth it! Do not forget that life is made of many wonderful sides, do not stick only to science, because anything else in your life can make you (and your science) way better. Take care of yourself and of your beloved people. Any minute of love that you could not enjoy with them is a wasted piece of life.

But I also know that young Anna would make her own decisions… as it should be!!!

An outing in nature with calçotada, Aiguamúrcia, mas d'en Ferran. Rodrigo, Víctor, Adrián, Sara, Lorenzo, Sanae, Anna, Alba, Óscar, and Raimundo.