

**Interreg**



Cofinanciado por  
la Unión Europea  
Cofinanciado pela  
União Europeia

España – Portugal



## **Desarrollo de modelos predictivos mediante tecnología NIRS**

*Tecnologías avanzadas, innovadoras y  
digitales para el sector agroalimentario  
de la EUROACE*

*Évora, 18<sup>th</sup> and 19<sup>th</sup>, novembre, 2024*



Researcher of the Meat Quality Area of CICYTEX

PhD in Food Science

[alberto.ortiz@juntaex.es](mailto:alberto.ortiz@juntaex.es) 



Degree in food science and technology

Meat Quality Area of CICYTEX

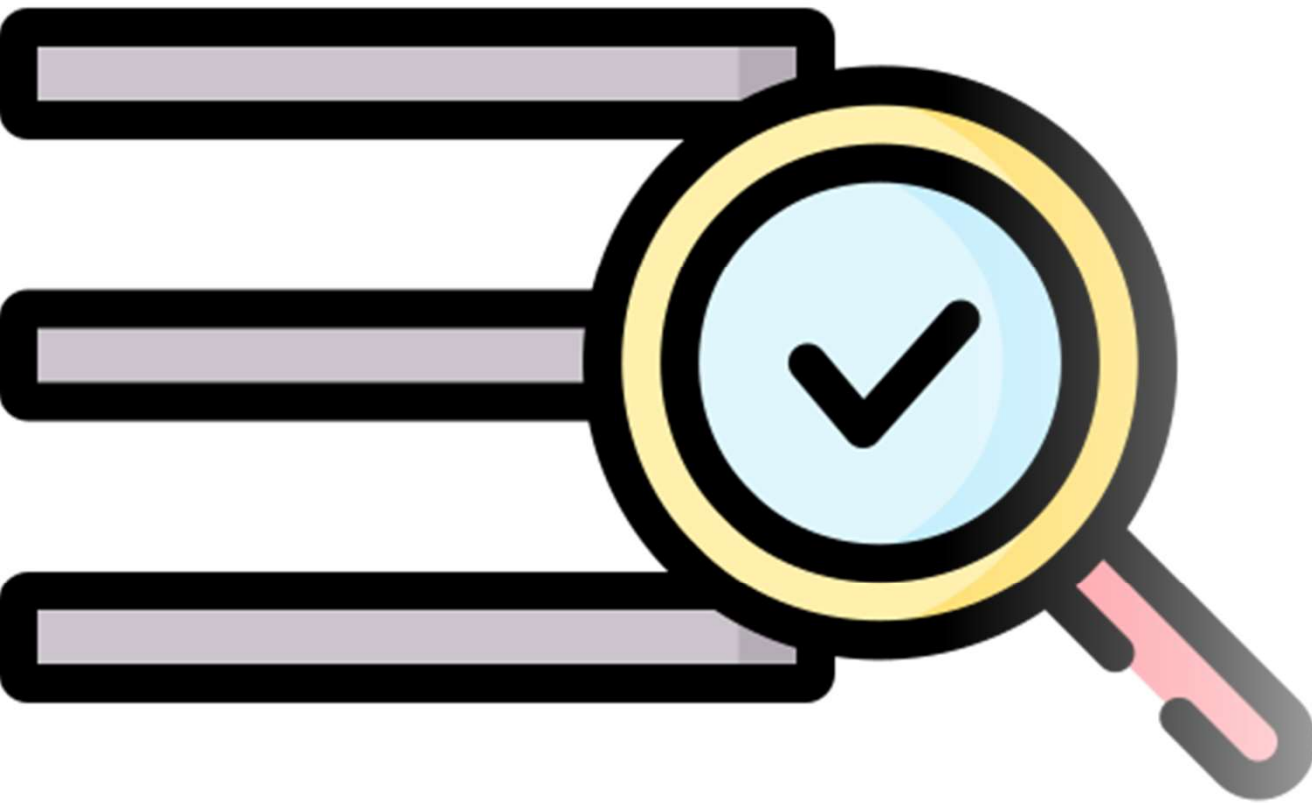
[lucia.leon@juntaex.es](mailto:lucia.leon@juntaex.es) 



**Centro de Investigaciones Científicas y  
Tecnológicas de Extremadura, CICYTEX.**

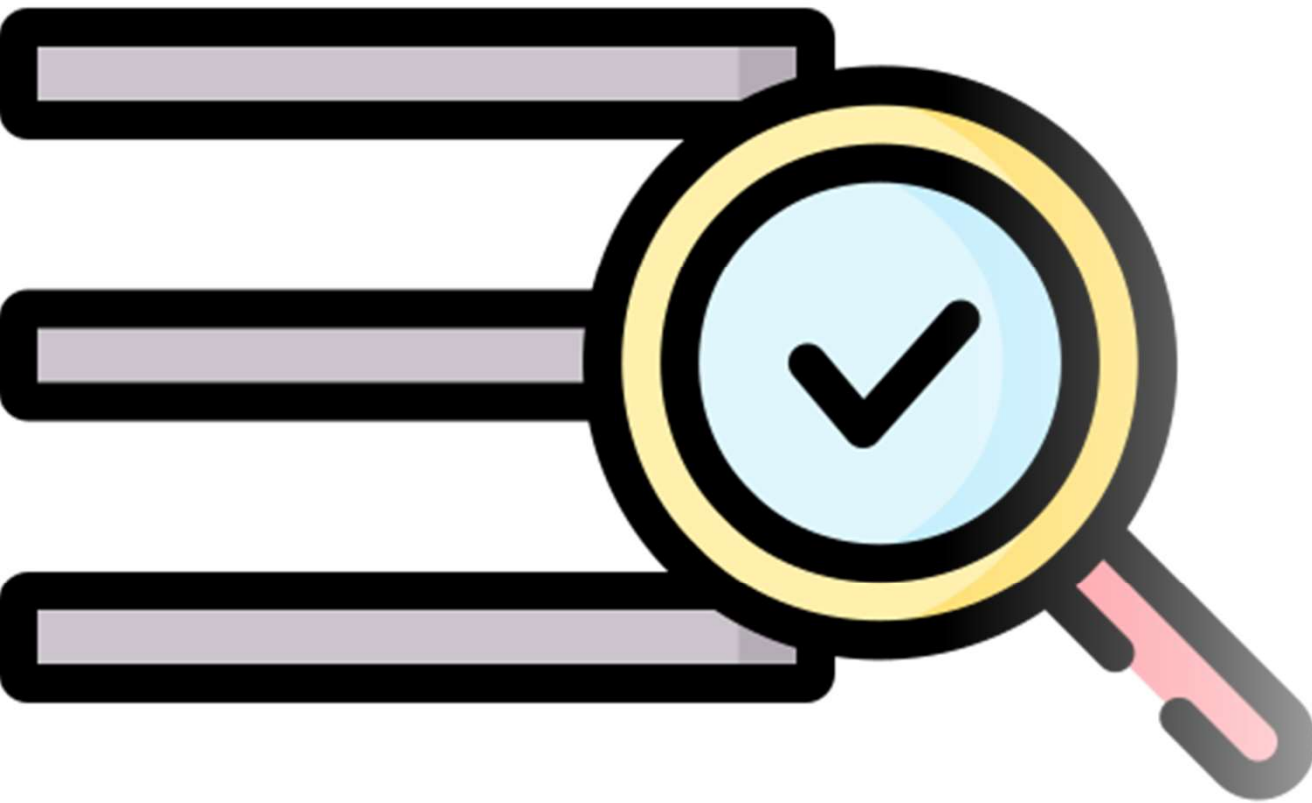
A5, km 472. 06187 Guadajira (Badajoz)





## Index

- What is NIRS technology?
- The Unscrambler
  - The construction of the matrix
  - Quantitative models
  - Qualitative models
  - Pre-treatments



## Index

- **What is NIRS technology?**
- The Unscrambler
  - The construction of the matrix
  - Quantitative models
  - Qualitative models
  - Pre-treatments





**What is NIRS technology?**



# What is NIRS technology?

**N**ear

**I**nfra**R**ed

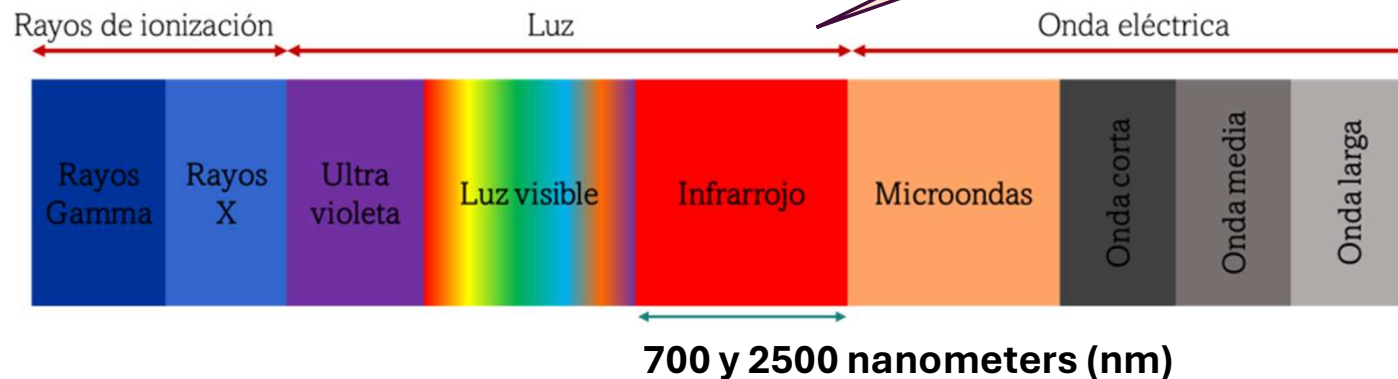
**S**pectroscopy



# What is NIRS technology?

## Near InfraRed Spectroscopy

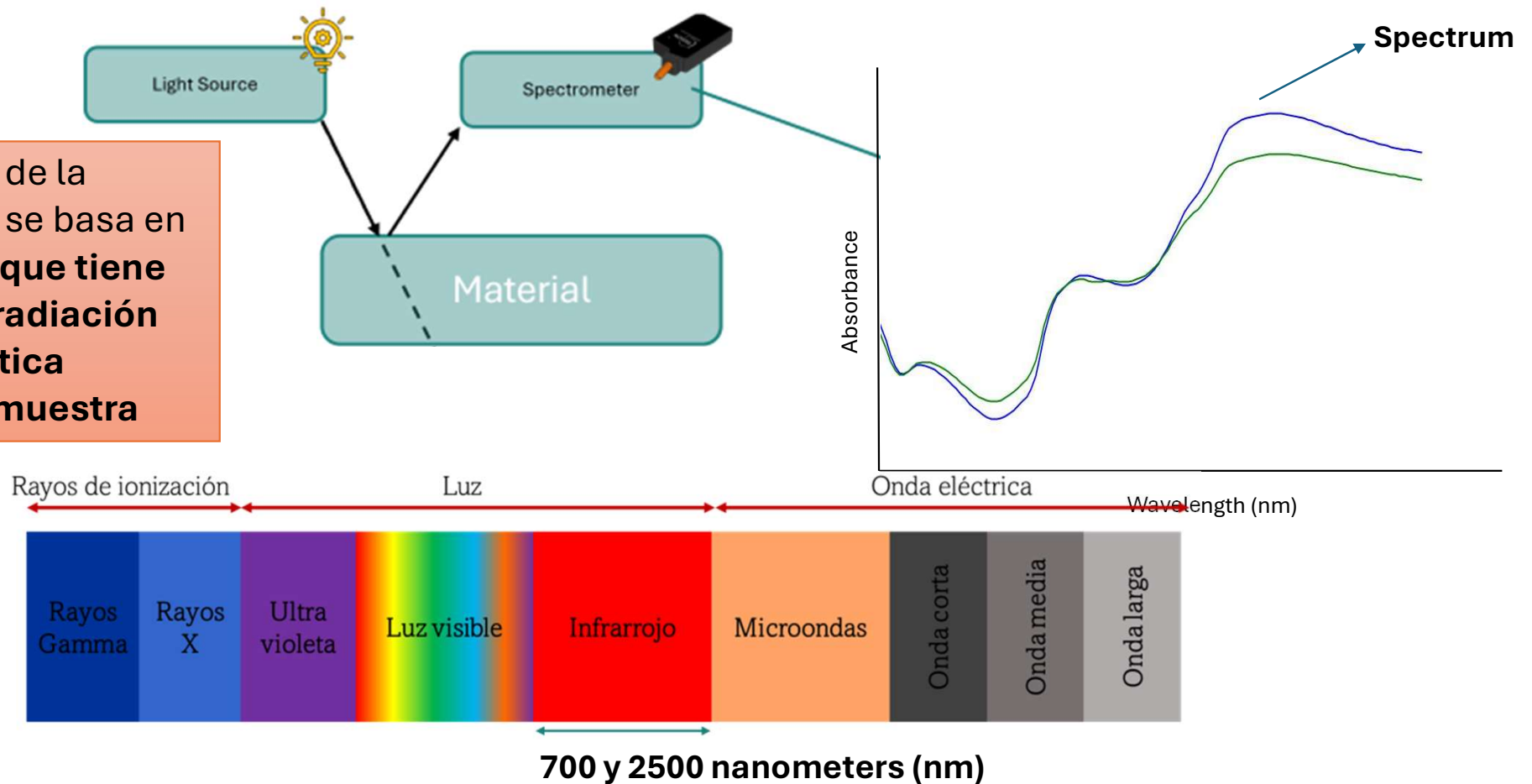
Near Infrared Spectroscopy (NIR) is an **analytical technique** that uses the **near infrared region** of the electromagnetic spectrum.

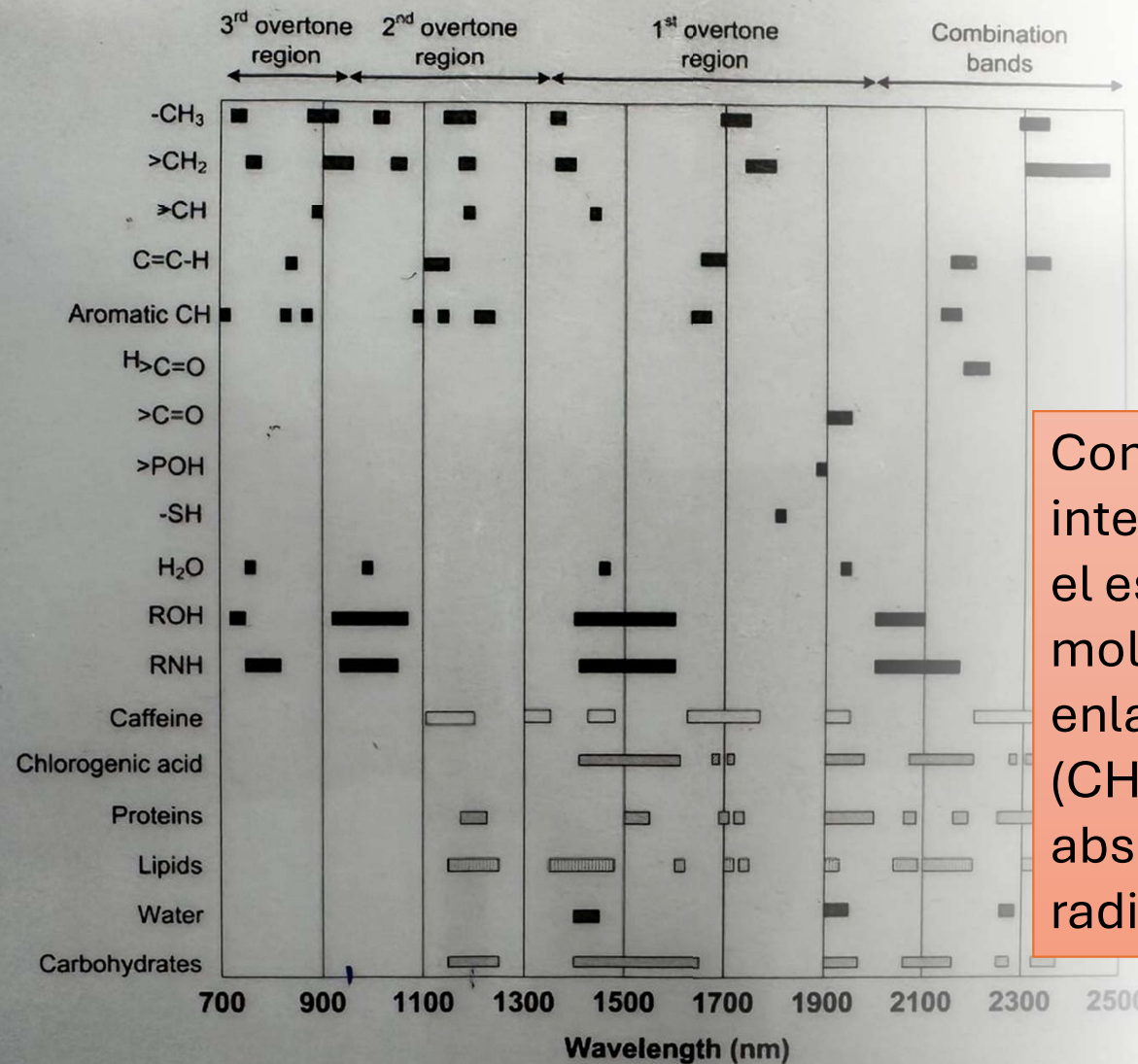




# What is NIRS technology?

El fundamento de la tecnología NIR se basa en la **interacción que tiene lugar entre la radiación electromagnética infrarroja y la muestra**





# What is NIRS technology?

Como consecuencia de esta interacción se producen cambios en el estado vibracional de las moléculas. Así, en función de los enlaces presentes en las moléculas (CH, NH, OH, CO) de la muestra, se absorberá una determinada radiación.

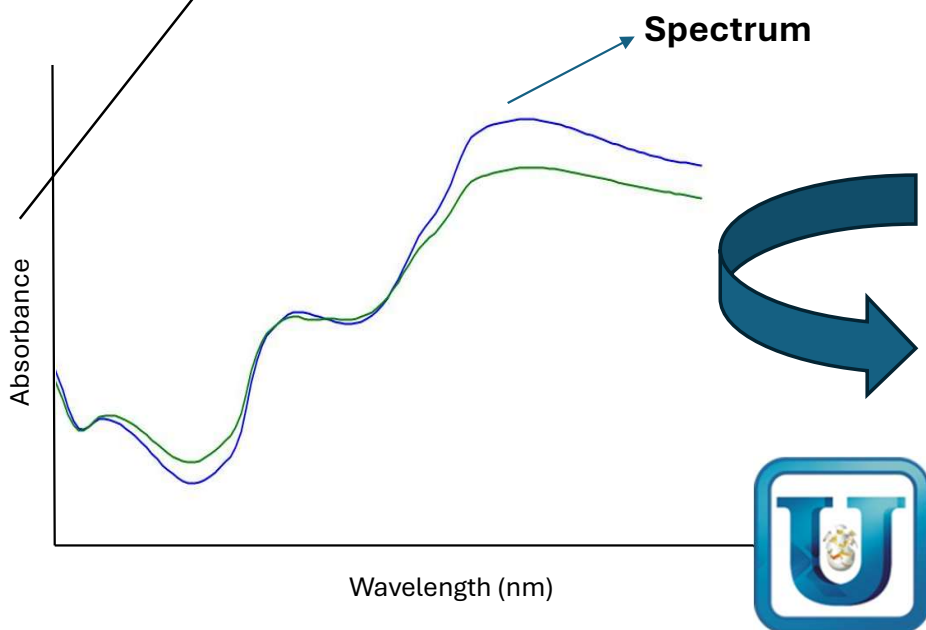
• Barbin et al., 2014



# What is NIRS technology?

Absorbancia =  $\text{Log}(1/R)$ , donde R son los valores de Reflectancia o cociente entre la radiación reflejada por la muestra y la incidente sobre la misma

Absorbance values



Espectros Uva NIRCROP 2023 [Espectros Uva NIRCROP 2023.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Data Matrix		908.1	914.294	920.489	926.683	932.877	939.072	945.266	951.46	957.655	963.849	970.044	976.238	982.432	988.626
		6	7	8	9	10	11	12	13	14	15	16	17	18	19
19 07 Temp	1	0,8915	0,8823	0,8723	0,8630	0,8559	0,8537	0,8555	0,8630	0,8703	0,8770	0,8785	0,8796	0,8793	0,8793
19 07 Temp	2	0,9024	0,8848	0,8686	0,8545	0,8430	0,8397	0,8435	0,8534	0,8652	0,8736	0,8767	0,8770	0,8786	0,8786
19 07 Temp	3	0,9053	0,8869	0,8694	0,8516	0,8378	0,8340	0,8376	0,8482	0,8596	0,8674	0,8721	0,8728	0,8715	0,8715
19 07 Temp	4	0,9207	0,9045	0,8884	0,8735	0,8627	0,8605	0,8622	0,8739	0,8847	0,8928	0,8962	0,8981	0,8978	0,8978
19 07 Temp	5	0,7748	0,7540	0,7323	0,7120	0,6974	0,6928	0,6958	0,7058	0,7188	0,7280	0,7313	0,7321	0,7309	0,7309
19 07 Temp	6	0,9609	0,9456	0,9307	0,9153	0,9053	0,9034	0,9079	0,9171	0,9281	0,9354	0,9391	0,9417	0,9416	0,9416
19 07 Temp	7	0,8418	0,8223	0,8051	0,7860	0,7721	0,7686	0,7699	0,7813	0,7925	0,8014	0,8044	0,8051	0,8051	0,8051
19 07 Temp	8	0,8708	0,8541	0,8374	0,8213	0,8094	0,8047	0,8083	0,8162	0,8279	0,8363	0,8405	0,8424	0,8419	0,8419
19 07 Temp	9	0,8761	0,8581	0,8418	0,8259	0,8123	0,8074	0,8095	0,8189	0,8305	0,8396	0,8428	0,8446	0,8439	0,8439
19 07 Temp	10	0,8154	0,7934	0,7716	0,7503	0,7355	0,7302	0,7335	0,7450	0,7575	0,7666	0,7712	0,7719	0,7713	0,7713
19 07 Temp	11	0,8253	0,8060	0,7865	0,7666	0,7524	0,7485	0,7505	0,7619	0,7742	0,7838	0,7881	0,7889	0,7876	0,7876
19 07 Temp	12	0,8998	0,8812	0,8640	0,8470	0,8325	0,8284	0,8303	0,8390	0,8477	0,8557	0,8584	0,8596	0,8581	0,8581
19 07 Temp	13	0,8898	0,8707	0,8521	0,8355	0,8231	0,8188	0,8235	0,8316	0,8449	0,8540	0,8579	0,8600	0,8594	0,8594
19 07 Temp	14	0,8168	0,7967	0,7773	0,7567	0,7412	0,7361	0,7362	0,7442	0,7528	0,7596	0,7627	0,7621	0,7617	0,7617
19 07 Temp	15	0,9015	0,8851	0,8677	0,8516	0,8379	0,8347	0,8372	0,8487	0,8591	0,8683	0,8727	0,8742	0,8731	0,8731
19 07 Temp	16	0,9169	0,9004	0,8856	0,8716	0,8601	0,8550	0,8546	0,8608	0,8683	0,8739	0,8761	0,8767	0,8765	0,8765
19 07 Temp	17	0,9249	0,9089	0,8956	0,8817	0,8693	0,8650	0,8652	0,8722	0,8798	0,8871	0,8890	0,8901	0,8896	0,8896
19 07 Temp	18	0,8227	0,8027	0,7820	0,7627	0,7486	0,7426	0,7449	0,7518	0,7611	0,7683	0,7708	0,7710	0,7694	0,7694
19 07 Temp	19	0,8860	0,8684	0,8521	0,8345	0,8201	0,8139	0,8141	0,8223	0,8305	0,8374	0,8408	0,8421	0,8410	0,8410
19 07 Temp	20	0,8472	0,8301	0,8127	0,7961	0,7821	0,7757	0,7766	0,7828	0,7911	0,7980	0,8014	0,8021	0,8008	0,8008
19 07 Temp	21	0,9218	0,9051	0,8897	0,8742	0,8607	0,8565	0,8588	0,8672	0,8766	0,8856	0,8888	0,8907	0,8893	0,8893
19 07 Temp	22	0,8924	0,8732	0,8544	0,8384	0,8258	0,8206	0,8237	0,8347	0,8470	0,8578	0,8623	0,8627	0,8616	0,8616
19 07 Temp	23	0,9649	0,9485	0,9331	0,9171	0,9048	0,9014	0,9036	0,9125	0,9214	0,9288	0,9327	0,9331	0,9333	0,9333
19 07 Temp	24	0,9285	0,9098	0,8927	0,8776	0,8661	0,8616	0,8646	0,8734	0,8850	0,8948	0,8980	0,9001	0,8994	0,8994
19 07 Temp	25	0,9410	0,9258	0,9120	0,8996	0,8882	0,8832	0,8832	0,8886	0,8941	0,8988	0,9023	0,9015	0,9019	0,9019
19 07 Temp	26	0,9596	0,9456	0,9314	0,9169	0,9081	0,9032	0,9052	0,9141	0,9230	0,9300	0,9334	0,9356	0,9346	0,9346
19 07 Temp	27	0,9251	0,9115	0,8978	0,8848	0,8748	0,8729	0,8736	0,8806	0,8880	0,8947	0,8977	0,8991	0,8995	0,8995
19 07 Temp	28	0,9372	0,9210	0,9068	0,8931	0,8836	0,8798	0,8837	0,8925	0,9019	0,9101	0,9137	0,9161	0,9154	0,9154
19 07 Temp	29	0,8702	0,8510	0,8314	0,8135	0,7994	0,7931	0,7927	0,7992	0,8082	0,8145	0,8171	0,8169	0,8169	0,8169

Name: Data Matrix  
Size: 780X135  
Created: 15/04/2024 12:...

The Unscrambler

CAMO software for analytical modeling and prediction

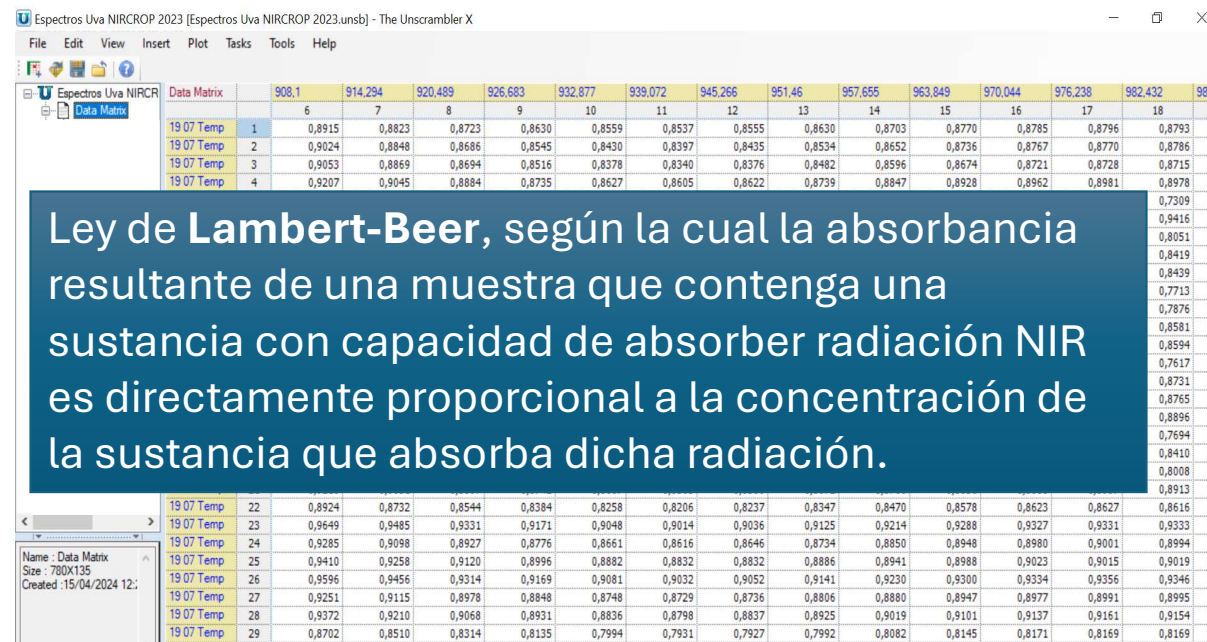


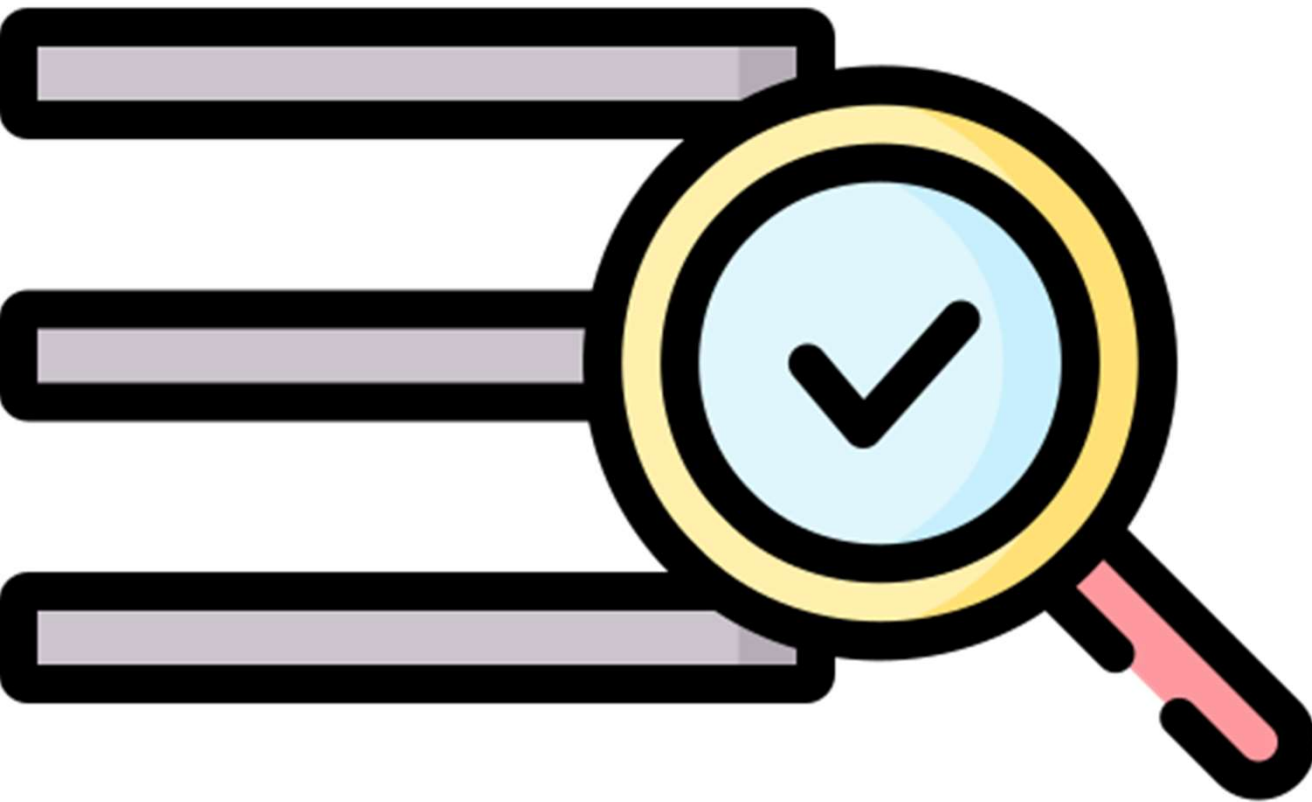


### Absorbance values



## CAMO software for analytical modeling and prediction





## Index

- What is NIRS technology?
- **The Unscrambler**
  - The construction of the matrix
  - Quantitative models
  - Qualitative models
  - Pre-treatments

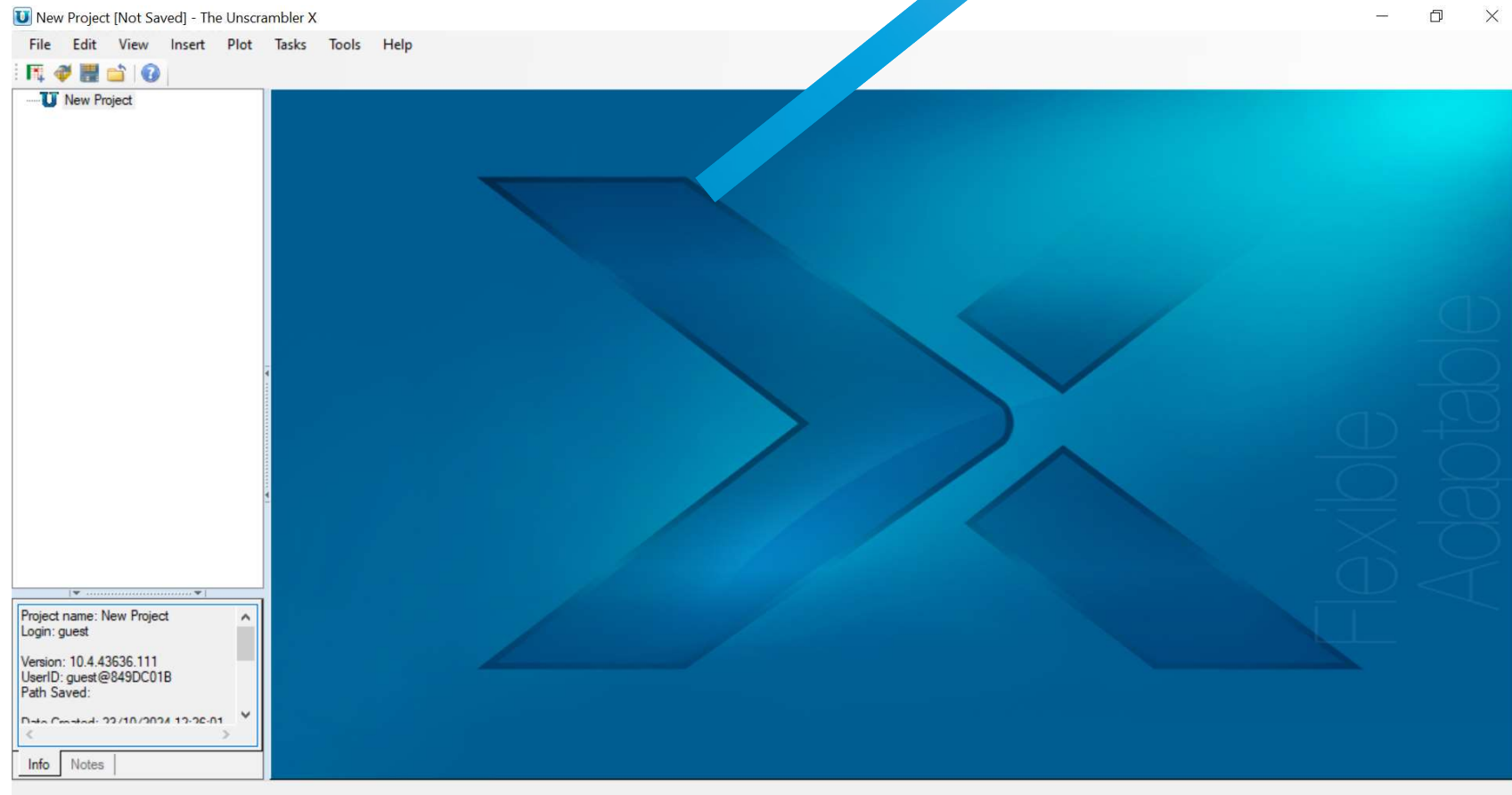


# The Unscrambler



## Workspace

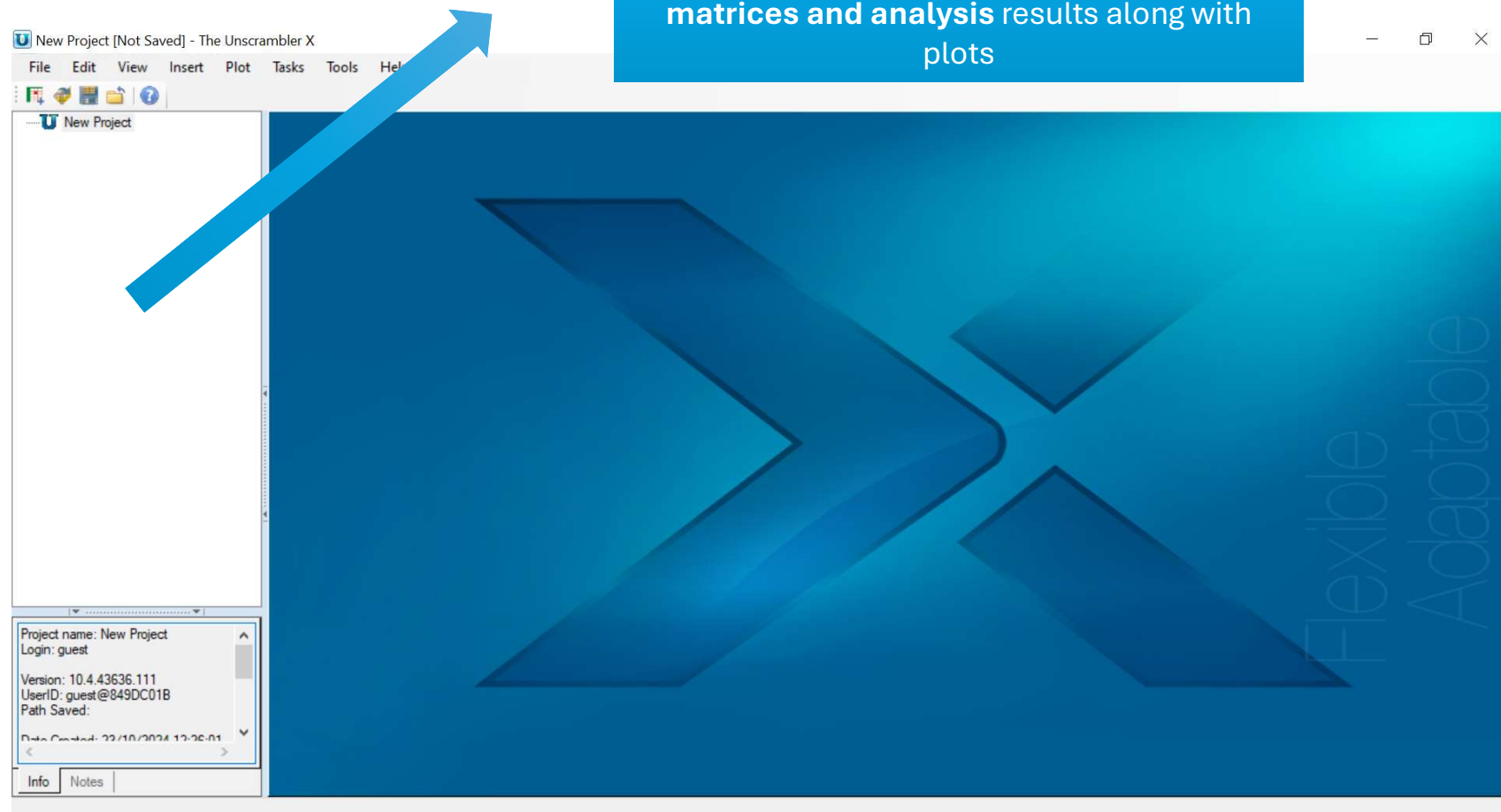
The **Workspace** occupies the largest area of the application window. It is a **Viewer** which displays results either graphically as plots or numerically as tables.



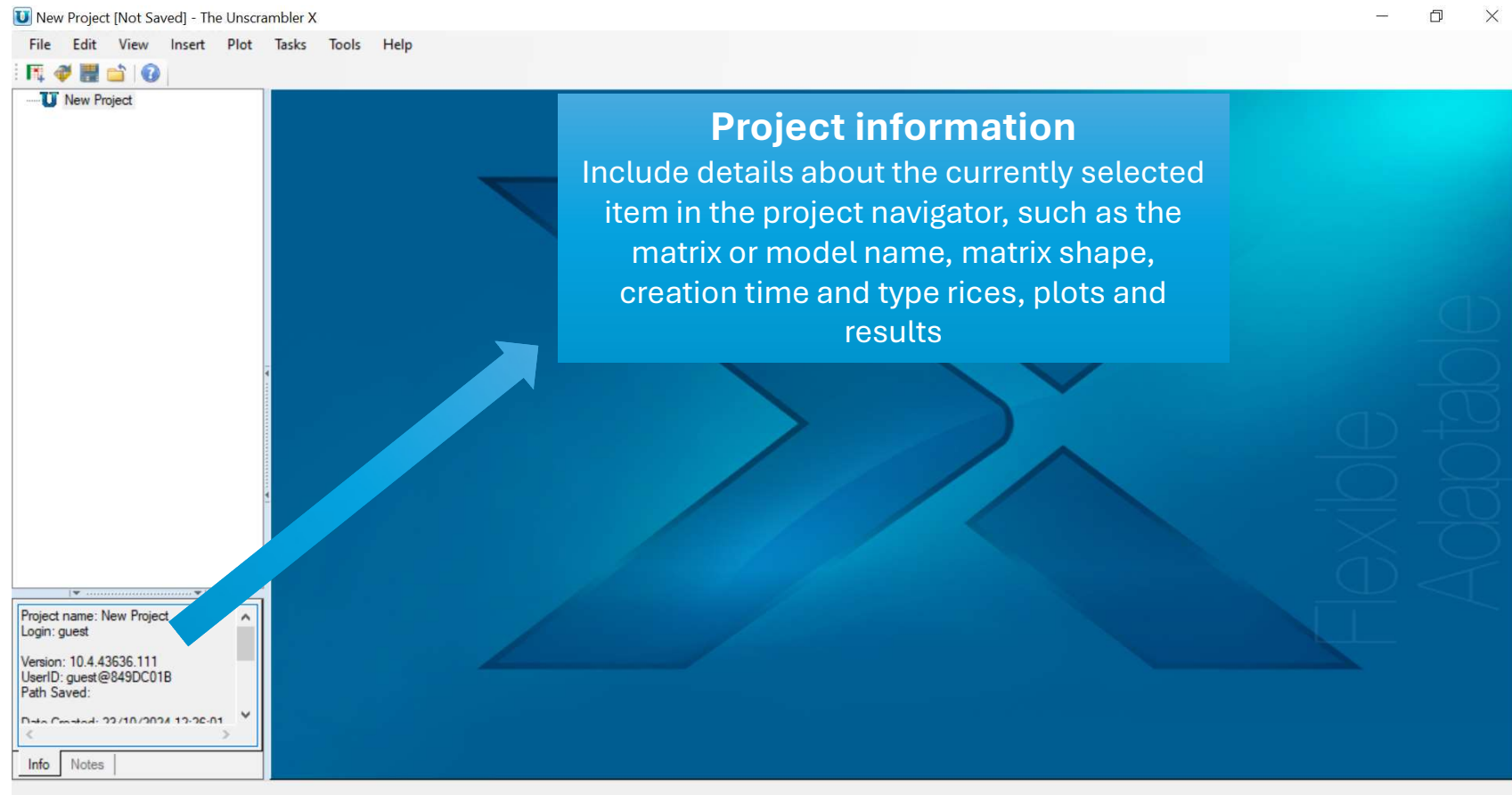
# The Unscrambler

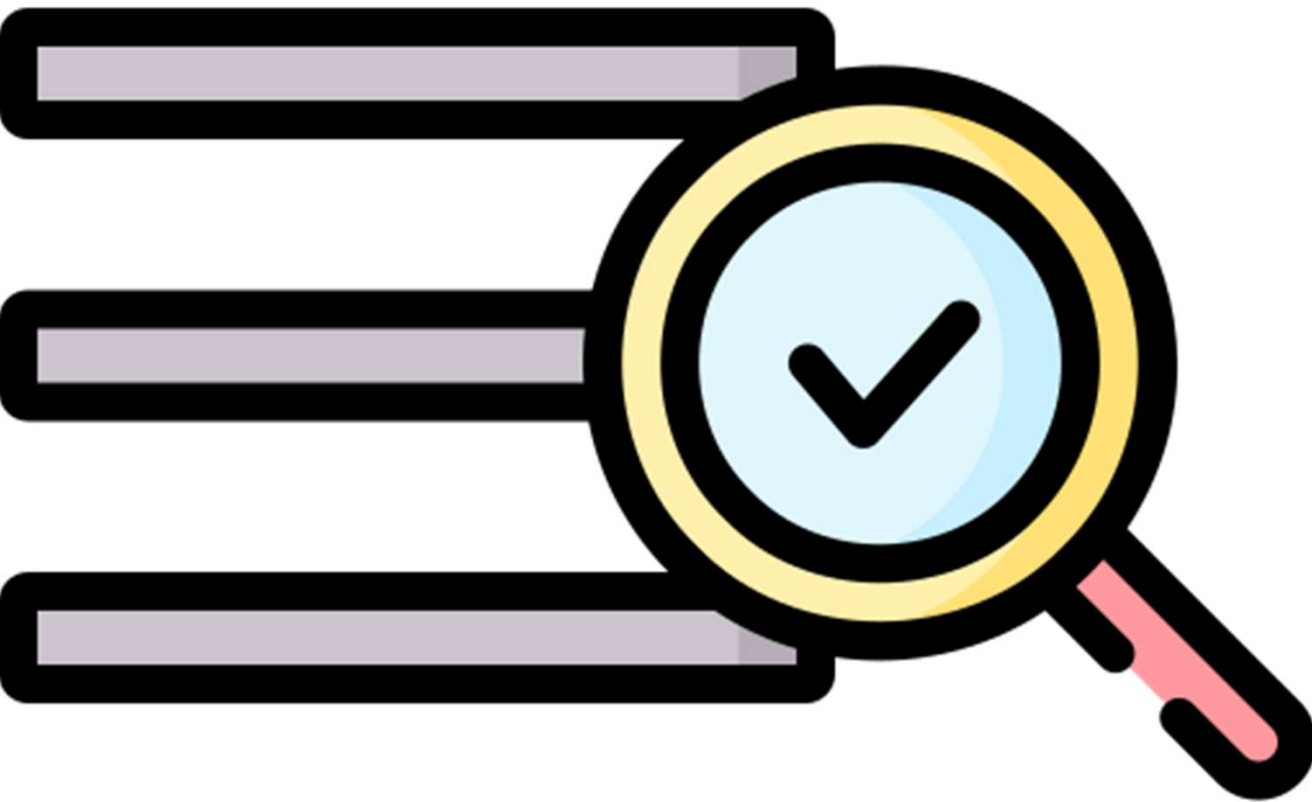


**Project navigator**  
Is a **tree-like structure** consisting of **data matrices** and **analysis results** along with **plots**



# The Unscrambler





## Index

- What is NIRS technology?
- The Unscrambler
  - **The construction of the matrix**
  - Quantitative models
  - Qualitative models
  - Pre-treatments

# The construction of the matrix

- Data import
- Reduce (average)
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Plot line

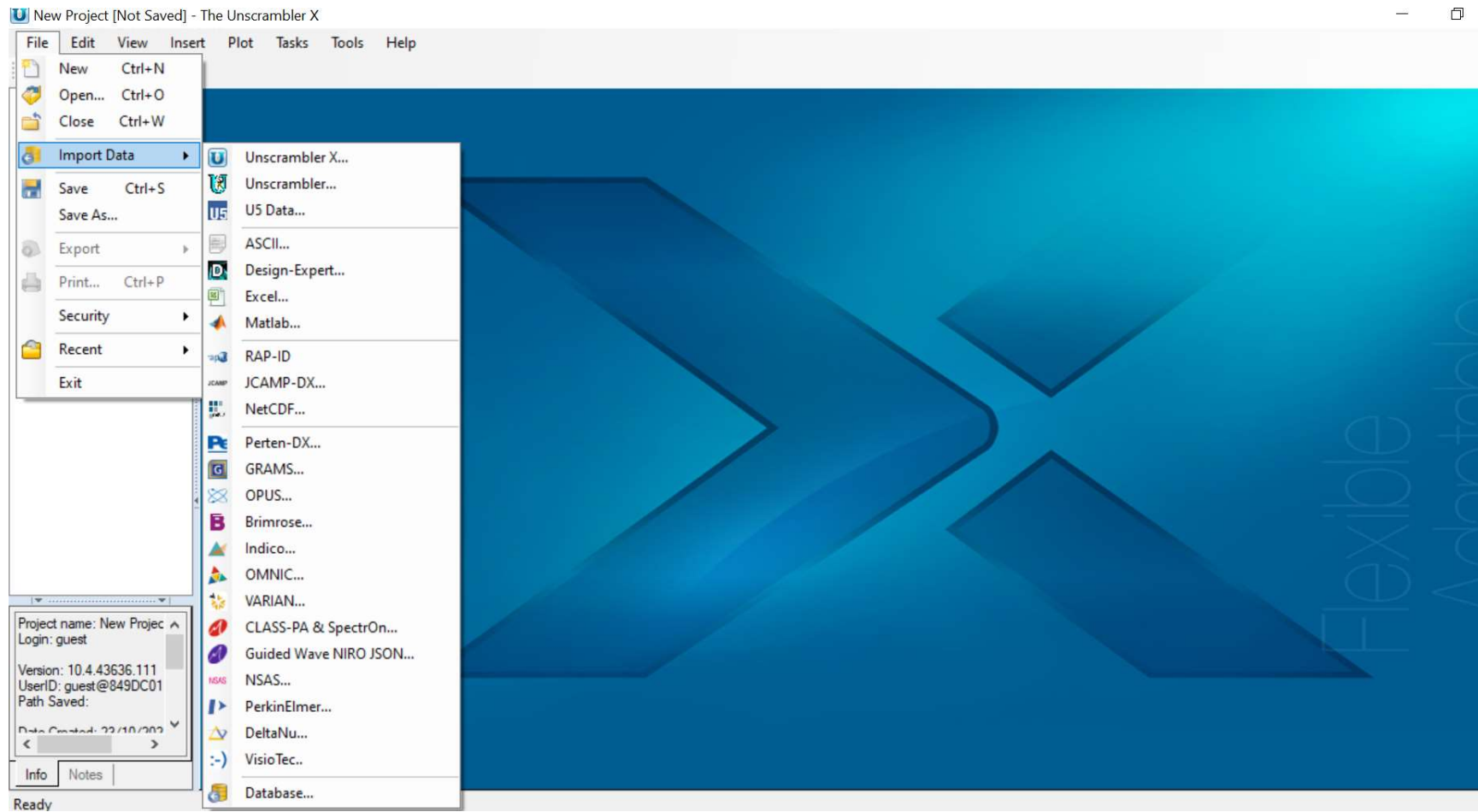


# The construction of the matrix

- **Data import**
- Reduce (average)
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Plot line



- Data import





- Data import

Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Data matrix oleico.unsb] - The Unscrambler X

Matriz para trabajar													97
		908.1	914.294	920.489	926.683	932.877	939.072	945.266	951.46	957.655	963.849	970.044	
		18	19	20	21	22	23	24	25	26	27	28	
1012127259-6947-cal-1.sam	7	0,5775	0,5645	0,5550	0,5434	0,5251	0,4994	0,4750	0,4582	0,4488	0,4431	0,4393	
1012128416-6950-Cal-1.sam	8	0,5927	0,5810	0,5719	0,5613	0,5440	0,5190	0,4949	0,4782	0,4686	0,4631	0,4594	
1012214133-6950-Cal-1.sam	9	0,6651	0,6530	0,6428	0,6311	0,6122	0,5833	0,5554	0,5362	0,5246	0,5176	0,5128	
1012214379-6947-cal-1.sam	10	0,5593	0,5463	0,5362	0,5240	0,5052	0,4784	0,4532	0,4360	0,4261	0,4204	0,4165	
1012214713-6951-Cal-1.sam	11	0,6161	0,6040	0,5952	0,5838	0,5645	0,5369	0,5105	0,4918	0,4816	0,4747	0,4702	
1012215710-6946-cal-1.sam	12	0,5091	0,4957	0,4863	0,4750	0,4569	0,4320	0,4080	0,3920	0,3832	0,3778	0,3745	
1012213778-6944-cal-1.sam	13	0,6615	0,6486	0,6392	0,6276	0,6071	0,5780	0,5497	0,5299	0,5187	0,5116	0,5075	
1012128836-6944-cal-1.sam	14	0,6287	0,6162	0,6067	0,5943	0,5740	0,5462	0,5196	0,5004	0,4895	0,4831	0,4788	
1012129079-6944-cal-1.sam	15	0,5009	0,4884	0,4784	0,4672	0,4500	0,4248	0,4010	0,3848	0,3760	0,3710	0,3672	
1012125453-6947-cal-1.sam	16	0,5496	0,5367	0,5270	0,5157	0,4978	0,4724	0,4480	0,4316	0,4226	0,4174	0,4141	
1012127792-6947-cal-1.sam	17	0,5368	0,5236	0,5136	0,5018	0,4828	0,4564	0,4315	0,4147	0,4050	0,3994	0,3957	
1012214072-6944-cal-1.sam	18	0,6025	0,5888	0,5790	0,5665	0,5464	0,5193	0,4929	0,4748	0,4645	0,4585	0,4547	
1012213938-6948-Cal-1.sam	19	0,5965	0,5829	0,5735	0,5620	0,5418	0,5134	0,4853	0,4669	0,4562	0,4497	0,4454	
1012214478-6949-Cal-1.sam	20	0,5972	0,5851	0,5762	0,5652	0,5462	0,5197	0,4946	0,4775	0,4678	0,4618	0,4579	
1012214058-6944-cal-1.sam	21	0,6305	0,6163	0,6063	0,5944	0,5738	0,5463	0,5196	0,5013	0,4906	0,4843	0,4800	
1012214195-6944-cal-1.sam	22	0,6334	0,6199	0,6107	0,5987	0,5780	0,5487	0,5208	0,5015	0,4903	0,4836	0,4790	
1012127952-6944-cal-1.sam	23	0,4929	0,4804	0,4707	0,4597	0,4421	0,4165	0,3919	0,3755	0,3666	0,3614	0,3576	
1012214416-6950-Cal-1.sam	24	0,5992	0,5862	0,5763	0,5647	0,5460	0,5183	0,4921	0,4742	0,4639	0,4577	0,4537	
1012214355-6944-cal-1.sam(1)	25	0,6017	0,5882	0,5781	0,5650	0,5447	0,5168	0,4901	0,4719	0,4616	0,4555	0,4510	
1012125477-6949-Cal-1.sam	26	0,6298	0,6198	0,6114	0,6002	0,5812	0,5531	0,5257	0,5068	0,4959	0,4894	0,4849	
1012127617-6950-Cal-1.sam	27	0,6167	0,6038	0,5943	0,5831	0,5638	0,5363	0,5093	0,4912	0,4807	0,4745	0,4703	
1012125293-6944-cal-1.sam	28	0,5888	0,5783	0,5694	0,5596	0,5444	0,5222	0,5007	0,4861	0,4781	0,4733	0,4702	
101224850-6948-Cal-1.sam	29	0,5945	0,5819	0,5718	0,5601	0,5420	0,5145	0,4880	0,4705	0,4603	0,4540	0,4501	
1012215154-6949-Cal-1.sam	30	0,5669	0,5577	0,5498	0,5391	0,5219	0,4963	0,4712	0,4541	0,4443	0,4384	0,4347	
1012129390-6944-cal-1.sam	31	0,5157	0,5036	0,4934	0,4824	0,4655	0,4419	0,4196	0,4049	0,3972	0,3924	0,3887	
1012126658-6944-cal-1.sam	32	0,5358	0,5255	0,5174	0,5080	0,4930	0,4704	0,4490	0,4345	0,4266	0,4222	0,4189	
1012214874-6975-cal-1.sam	33	0,6086	0,5955	0,5854	0,5731	0,5531	0,5231	0,4942	0,4744	0,4631	0,4564	0,4520	
1012215215-6972-cal-1.sam	34	0,5336	0,5196	0,5094	0,4976	0,4784	0,4505	0,4233	0,4049	0,3941	0,3880	0,3838	
1012127730-6968-cal-1.sam	35	0,5620	0,5517	0,5433	0,5331	0,5149	0,4897	0,4648	0,4476	0,4380	0,4320	0,4280	
1012214690-6966-cal-1.sam	36	0,5897	0,5769	0,5676	0,5564	0,5369	0,5092	0,4835	0,4657	0,4553	0,4494	0,4449	



# The construction of the matrix

- Data import
- **Reduce (average)**
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Plot line



- Reduce (Average)

TE.unsb] - The Unscrambler X

Tasks Tools Help

[illegible]

It consists of an average of spectra. Since noise is considered as a random instrumental variation, averaging over samples (in case of replicates) may have, depending on the context, the following advantages:

- ☐ Increase precision
- ☐ Get more stable results
- ☐ Reduce noise
- ☐ Interpret the results more easily

- Reduce (Average)

NIR CROP TOMATE [NIR CROP TOMATE.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

NIR CROP TOMATE

- Data Matrix Original limpia
- Data Matrix Para trabajar\_Reduc
- Columnset
- Sello de tiempo
- Espectros
- Número de serie del equipo
- Temperatura
- Notas
- BRIX
- Acidez
- Licopeno
- Textura
- Line Plot
- PLS Brx Abs
- PLS Acidez Abs
- PLS Licopeno Abs
- PLS TXT Abs
- Data Matrix Para trabajar\_Reduc
- PLS SNV DE Brx
- PLS Acidez SNV DE
- PLS Licopeno SNV DE
- PLS SNV DE Textura
- Data Matrix Original limpia(1)
- Columnset

Data Matrix Original limpia(1)

	Fecha	pH	BRIX	Acidez (%)	Licopeno (m)	Textura	908,1	914,294	920,489	926,683	932,877	938,071
04 07 PARCELA 01 PLANTA 01-1.sam	1	04-07-2023					0,5433	0,5206	0,4992	0,4730	0,4454	0,4178
04 07 PARCELA 01 PLANTA 01-2.sam	2	04-07-2023					0,5429	0,5243	0,5045	0,4904	0,4873	0,4842
04 07 PARCELA 01 PLANTA 01-3.sam	3	04-07-2023					0,5715	0,5518	0,5325	0,5142	0,5037	0,4932
04 07 PARCELA 01 PLANTA 01-4.sam	4	04-07-2023					0,5884	0,5679	0,5488	0,5276	0,5092	0,4908
04 07 PARCELA 01 PLANTA 01-5.sam	5	04-07-2023					0,5461	0,5271	0,5047	0,4956	0,5048	0,4957
04 07 PARCELA 01 PLANTA 02-1.sam	6	04-07-2023					0,5775	0,5560	0,5354	0,5161	0,5034	0,4907
04 07 PARCELA 01 PLANTA 02-2.sam	7	04-07-2023					0,5614	0,5407	0,5207	0,4998	0,4843	0,4688
04 07 PARCELA 01 PLANTA 02-3.sam	8	04-07-2023					0,5840	0,5628	0,5427	0,5199	0,4991	0,4783
04 07 PARCELA 01 PLANTA 02-4.sam	9	04-07-2023					0,5716	0,5517	0,5327	0,5101	0,4890	0,4680
04 07 PARCELA 01 PLANTA 02-5.sam	10	04-07-2023					0,5939	0,5733	0,5539	0,5321	0,5105	0,4890
04 07 PARCELA 01 PLANTA 03-1.sam	11	04-07-2023					0,5695	0,5522	0,5338	0,5201	0,5162	0,5123
04 07 PARCELA 01 PLANTA 03-2.sam	12	04-07-2023					0,5283	0,5071	0,4867	0,4640	0,4441	0,4242
04 07 PARCELA 01 PLANTA 03-3.sam	13	04-07-2023					0,6188	0,6016	0,5854	0,5647	0,5441	0,5235
04 07 PARCELA 01 PLANTA 03-4.sam	14	04-07-2023					0,6260	0,6107	0,5961	0,5768	0,5575	0,5382
04 07 PARCELA 01 PLANTA 03-5.sam	15	04-07-2023					0,6283	0,6124	0,5971	0,5797	0,5639	0,5481
04 07 PARCELA 01 PLANTA 04-1.sam	16	04-07-2023					0,5670	0,5499	0,5332	0,5142	0,4972	0,4802
04 07 PARCELA 01 PLANTA 04-2.sam	17	04-07-2023					0,5733	0,5563	0,5407	0,5195	0,4972	0,4750
04 07 PARCELA 01 PLANTA 04-3.sam	18	04-07-2023					0,5453	0,5289	0,5139	0,4950	0,4752	0,4554
04 07 PARCELA 01 PLANTA 04-4.sam	19	04-07-2023					0,5696	0,5570	0,5420	0,5357	0,5435	0,5313
04 07 PARCELA 01 PLANTA 04-5.sam	20	04-07-2023					0,5122	0,4983	0,4823	0,4758	0,4865	0,4772
04 07 PARCELA 01 PLANTA 05-1.sam	21	04-07-2023					0,5345	0,5215	0,5075	0,4961	0,4924	0,4887
04 07 PARCELA 01 PLANTA 05-2.sam	22	04-07-2023					0,4906	0,4778	0,4623	0,4566	0,4675	0,4584
04 07 PARCELA 01 PLANTA 05-3.sam	23	04-07-2023					0,4850	0,4720	0,4561	0,4519	0,4673	0,4582
04 07 PARCELA 01 PLANTA 05-4.sam	24	04-07-2023					0,5376	0,5259	0,5145	0,5021	0,4908	0,4795
04 07 PARCELA 01 PLANTA 05-5.sam	25	04-07-2023					0,5150	0,5040	0,4936	0,4812	0,4720	0,4628
04 07 PARCELA 01 PLANTA 05-6.sam	26	04-07-2023					0,5453	0,5358	0,5276	0,5151	0,5005	0,4859
04 07 PARCELA 01 PLANTA 05-7.sam	27	04-07-2023					0,5182	0,5091	0,5003	0,4894	0,4778	0,4662
04 07 PARCELA 01 PLANTA 05-8.sam	28	04-07-2023					0,4781	0,4688	0,4589	0,4517	0,4501	0,4485
04 07 PARCELA 01 PLANTA 05-9.sam	29	04-07-2023					0,4879	0,4793	0,4697	0,4628	0,4610	0,4594
04 07 PARCELA 01 PLANTA 05-10.sam	30	04-07-2023					0,5130	0,5059	0,4991	0,4909	0,4851	0,4814
04 07 PARCELA 01 PLANTA 05-11.sam	31	04-07-2023					0,5213	0,5143	0,5072	0,4992	0,4914	0,4836
04 07 PARCELA 01 PLANTA 05-12.sam	32	04-07-2023					0,4949	0,4878	0,4802	0,4741	0,4704	0,4667
04 07 PARCELA 01 PLANTA 05-13.sam	33	04-07-2023					0,4550	0,4472	0,4384	0,4337	0,4353	0,4369
04 07 PARCELA 01 PLANTA 05-14.sam	34	04-07-2023					0,4931	0,4875	0,4817	0,4765	0,4744	0,4723
04 07 PARCELA 01 PLANTA 05-15.sam	35	04-07-2023					0,4639	0,4570	0,4495	0,4438	0,4424	0,4410
04 07 PARCELA 01 PLANTA 05-16.sam	36	04-07-2023					0,6189	0,6020	0,5853	0,5696	0,5611	0,5526
04 07 PARCELA 01 PLANTA 05-17.sam	37	04-07-2023					0,5371	0,5189	0,5007	0,4843	0,4760	0,4677
04 07 PARCELA 01 PLANTA 05-18.sam	38	04-07-2023					0,5595	0,5399	0,5190	0,5018	0,4973	0,4928
04 07 PARCELA 01 PLANTA 05-19.sam	39	04-07-2023					0,6153	0,5977	0,5801	0,5603	0,5439	0,5275
04 07 PARCELA 01 PLANTA 05-20.sam	40	04-07-2023					0,5471	0,5268	0,5071	0,4842	0,4642	0,4442
04 07 PARCELA 01 PLANTA 05-21.sam	41	04-07-2023					0,4480	0,4252	0,4015	0,3839	0,3795	0,3751
04 07 PARCELA 01 PLANTA 05-22.sam	42	04-07-2023					0,5161	0,4975	0,4753	0,4692	0,4958	0,4914
04 07 PARCELA 01 PLANTA 05-23.sam	43	04-07-2023					0,4751	0,4571	0,4360	0,4310	0,4590	0,4546
04 07 PARCELA 01 PLANTA 05-24.sam	44	04-07-2023					0,4656	0,4453	0,4224	0,4153	0,4395	0,4351
04 07 PARCELA 01 PLANTA 05-25.sam	45	04-07-2023					0,5813	0,5643	0,5450	0,5344	0,5346	0,5348
04 07 PARCELA 01 PLANTA 05-26.sam	46	04-07-2023					0,5898	0,5694	0,5507	0,5278	0,5063	0,4848
04 07 PARCELA 01 PLANTA 05-27.sam	47	04-07-2023					0,3695	0,3490	0,3247	0,3224	0,3705	0,3661
04 07 PARCELA 01 PLANTA 05-28.sam	48	04-07-2023					0,4461	0,4273	0,4050	0,4045	0,4555	0,4511
04 07 PARCELA 01 PLANTA 05-29.sam	49	04-07-2023					0,5390	0,5192	0,4999	0,4807	0,4672	0,4537
04 07 PARCELA 01 PLANTA 05-30.sam	50	04-07-2023					0,5408	0,5229	0,5049	0,4892	0,4809	0,4726
04 07 PARCELA 01 PLANTA 05-31.sam	51	04-07-2023					0,5729	0,5549	0,5370	0,5199	0,5088	0,4977

Transform

- Plot
- Export
- Range
- Duplicate Matrix...
- Rename
- Delete
- Spectra
- Save Matrix

Smoothing

- Normalize...
- Derivative
- Baseline...
- SNV...
- Spectroscopic...
- De-trending...
- MSC/EMSC...
- OSC...
- Deresolve...
- Reduce (Average)...
- Transpose
- Noise...
- Interpolate...
- Interaction and Square Effects...
- Weights...
- Center and Scale...
- Quantile Normalize...
- Compute General...
- Fill Missing...
- Correlation Optimized Warping...
- Piecewise Direct Standardization...
- Sample alignment...

Name: Data Matrix Original limpia(1)  
Size: 1490 X 136  
Created: 06/11/2024 8:50:12

Info Notes

PLS SNV DE Text... Data Matrix Ori... Data Matrix Ori...

Reduces the matrix size by averaging samples or variables



- Reduce (Average)

0,5283

### Reduce (Average)

?

X

Data Matrix: Data Matrix Original limpia(1) {1490x136}

Rows: All Cols: Espectros {125} Define...

Reduce along:

☐ Variables

☒ Samples

Settings:

☒ Reduction factor: 5

☐ Sample index

☐ Insert number of replicates

Sample index:

Matrix: Rows: Cols: Define...

OK Cancel

Data Matrix Original limpia(1)\_Reduced

		908,1	914,294	920,489
		1	2	3
04 07 PARCELA 01 PLANTA 01-3.sam	1	0,5585	0,5383	0,5179
04 07 PARCELA 01 PLANTA 02-3.sam	2	0,5777	0,5569	0,5371
04 07 PARCELA 01 PLANTA 03-3.sam	3	0,5942	0,5768	0,5598
04 07 PARCELA 01 PLANTA 04-3.sam	4	0,5535	0,5381	0,5224
04 07 PARCELA 01 PLANTA 05-3.sam	5	0,5125	0,5003	0,4868
04 07 PARCELA 01 PLANTA 06-3.sam	6	0,5085	0,4998	0,4911
04 07 PARCELA 01 PLANTA 07-3.sam	7	0,4856	0,4788	0,4714
04 07 PARCELA 01 PLANTA 08-3.sam	8	0,5756	0,5570	0,5384
04 07 PARCELA 01 PLANTA 09-3.sam	9	0,4972	0,4779	0,4560
04 07 PARCELA 01 PLANTA 10-3.sam	10	0,4970	0,4776	0,4570
04 07 PARCELA 01 PLANTA 11-3.sam	11	0,5127	0,4938	0,4739
04 07 PARCELA 01 PLANTA 12-3.sam	12	0,4961	0,4765	0,4554

0,4639

# The construction of the matrix

- Data import
- Reduce (average)
- **Insert row/column (to incorporate new samples or physical-chemical variables to predict)**
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Plot line



- Insert row (to incorporate new samples)

Espectros Uva NIRCROP 2023 [Data Matrix uva.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Espectros Uva NIRCROP Data Matrix

	Fecha	Variedad	pH	BRIX	Acidez (%)	908,1	914,294	920,489	926,683	932,877	939,072	945,266	
	1	2	3	4	5	6	7	8	9	10	11	12	
19 07 Temp	1	19-07	Tempranillo	3,4200	18,1300	0,3960	0,8915	0,8823	0,8723	0,8630	0,8559	0,8537	0,8555
19 07 Temp	2	19-07	Tempranillo	3,2200	15,3800	0,3470	0,9024	0,8848	0,8686	0,8545	0,8430	0,8397	0,8435
19 07 Temp	3	19-07	Tempranillo	3,2500	13,7500	0,6760	0,9053	0,8869	0,8694	0,8516	0,8378	0,8340	0,8376
19 07 Temp	4	19-07	Tempranillo	3,2300	15,9800	0,3810	0,9207	0,9045	0,8884	0,8735	0,8627	0,8605	0,8622
19 07 Temp				2,6500	7,7300	1,0490	0,7748	0,7540	0,7323	0,7120	0,6974	0,6928	0,6958
19 07 Temp				3,2000	17,1900	0,3310	0,9609	0,9456	0,9307	0,9153	0,9053	0,9034	0,9079
19 07 Temp				2,8000	14,4500	0,6160	0,8418	0,8223	0,8051	0,7860	0,7721	0,7686	0,7699
19 07 Temp				2,8800	17,6600	0,7260	0,8708	0,8541	0,8374	0,8213	0,8094	0,8047	0,8083
19 07 Temp				3,0400	14,4500	0,5610	0,8761	0,8581	0,8418	0,8259	0,8123	0,8074	0,8095
19 07 Temp				2,8700	11,3700	0,4180	0,8154	0,7934	0,7716	0,7503	0,7355	0,7302	0,7335
19 07 Temp				2,6200	13,0100	0,8600	0,8253	0,8060	0,7865	0,7666	0,7524	0,7485	0,7505
19 07 Temp				3,1400	12,4700	0,5370	0,8998	0,8812	0,8640	0,8470	0,8325	0,8284	0,8303
19 07 Temp				2,9500	16,6700	0,3980	0,8888	0,8703	0,8534	0,8355	0,8231	0,8188	0,8235
19 07 Temp				2,8100	14,4800	0,5020	0,8888	0,8703	0,8534	0,8355	0,8231	0,8188	0,8235
19 07 Temp				2,9600	18,1500	0,2800	0,8379	0,8194	0,8015	0,7836	0,7657	0,7478	0,7299
19 07 Temp				2,8800	15,9300	0,2580	0,8601	0,8416	0,8237	0,8058	0,7879	0,7699	0,7520
19 07 Temp				2,7900	12,7000	0,6090	0,8693	0,8508	0,8329	0,8150	0,7971	0,7792	0,7613
19 07 Temp				2,8300	14,3800	1,2220	0,7486	0,7301	0,7122	0,6943	0,6764	0,6585	0,6406
19 07 Temp				2,8600	16,3700	0,5870	0,8201	0,8016	0,7837	0,7658	0,7479	0,7299	0,7120
19 07 Temp				2,5400	12,2400	0,7860	0,8472	0,8287	0,8108	0,7929	0,7750	0,7571	0,7392
19 07 Temp				3,3970	18,1300	0,3960	0,8915	0,8723	0,8534	0,8345	0,8156	0,7967	0,7778
19 07 Temp				3,5130	18,1300	0,3960	0,8915	0,8723	0,8534	0,8345	0,8156	0,7967	0,7778
19 07 Temp				3,5340	18,1300	0,3960	0,8915	0,8723	0,8534	0,8345	0,8156	0,7967	0,7778
19 07 Temp				3,2500	16,9700	0,2830	0,9285	0,9098	0,8927	0,8776	0,8661	0,8616	0,8646
19 07 Temp				3,0900	14,3000	0,4970	0,9410	0,9258	0,9120	0,8996	0,8882	0,8832	0,8832
19 07 Temp				3,0900	15,6600	0,4600	0,9596	0,9456	0,9314	0,9169	0,9081	0,9032	0,9052
19 07 Temp	27	19-07	Tempranillo	2,9800	15,0500	0,4130	0,9251	0,9115	0,8978	0,8848	0,8748	0,8729	0,8736
19 07 Temp	28	19-07	Tempranillo	2,9500	17,7700	0,3820	0,9372	0,9210	0,9068	0,8931	0,8836	0,8798	0,8837
19 07 Temp	29	19-07	Tempranillo	2,8600	13,6900	0,6130	0,8702	0,8510	0,8314	0,8135	0,7994	0,7931	0,7927
19 07 Temp	30	19-07	Tempranillo	3,0300	17,2100	0,3550	0,9194	0,9033	0,8869	0,8721	0,8597	0,8556	0,8550

Insert Rows

Number of new rows : 1

OK Cancel

Info Notes

Calibration Data Matrix

Name : Data Matrix  
Size : 780X135  
Created : 15/04/2024 12:...



- Insert column (to incorporate physical-chemical variables to predict)

Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Data matrix oleico.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Data matrix oleico.unsb] - The Unscrambler X

Matriz para trabajar

		C18:1 n-9	C18:2 n-6	C18:3 n-3	C20:0	C20:1	AGS	AGMI	AGPI	908.1	914.294	920.489	926.6
		10	11	12	13	14	15	16	17	18	19	20	
1012127259-6947-cal-1.sam	7	54,5900	9,0100	0,6600	0,2200	1,9700	31,1800	59,1500	9,6700	0,5775	0,5645	0,5550	
1012128416-6950-Cal-1.sam	8	58,4700	8,8600	0,5700	0,2100	1,9500	28,0700	62,5000	9,4300	0,5927	0,5810	0,5719	
1012214133-6950-Cal-1.sam	9	56,9300	8,3300	0,4900	0,2000	1,5800	30,2100	60,9800	8,8200	0,6651	0,6530	0,6428	
1012214379-6947-cal-1.sam	10	54,4000	9,5000	0,6500	0,2300	1,8600	30,7800	59,0700	10,1500	0,5593	0,5463	0,5362	
1012214713-6951-Cal-1.sam	11	57,5900	7,9400	0,8700	0,1800	1,7000	29,3900	61,8000	8,8100	0,6161	0,6040	0,5952	
1012215710-6946-cal-1.sam	12	53,5700	7,4000	0,5900	0,2200	1,5000	34,4900	57,5200	7,9900	0,5091	0,4957	0,4863	
1012213778-6944-cal-1.sam	13	59,3400	10,0300	1,0000	0,1800	1,6000	25,6700	63,3000	11,0300	0,6615	0,6486	0,6392	
1012128836-6944-cal-1.sam	14	59,1100	9,0400	0,7500	0,2200	1,7300	27,2500	62,9600	9,7900	0,6287	0,6162	0,6067	
1012129079-6944-cal-1.sam	15	51,8400	8,5900	0,9100	0,2100	1,7600	34,5400	55,9600	9,5000	0,5009	0,4884	0,4784	
1012125453-6947-cal-1.sam	16	53,8700	7,5700	0,5600	0,2700	2,0600	33,6300	58,2400	8,1300	0,5496	0,5367	0,5270	
1012127792-6947-cal-1.sam	17	53,6700	8,3200	0,7700	0,2200	1,5500	33,5200	57,4000	9,0800	0,5368	0,5236	0,5136	
1012214072-6944-cal-1.sam	18	59,1700	9,6800	0,7700	0,2200	1,9600	26,4900	63,0600	10,4500	0,6025	0,5888	0,5790	
1012213938-6948-Cal-1.sam	19	58,4100	10,5600	0,8400	0,2100	1,4900	26,5600	62,0300	11,4100	0,5965	0,5829	0,5735	
1012214478-6949-Cal-1.sam	20	54,8300	11,9200	0,8400	0,2000	1,6600	28,5200	58,7300	12,7500	0,5972	0,5851	0,5762	
1012214058-6944-cal-1.sam	21	58,0600	9,3600	0,9900	0,2400	1,6000	27,7500	61,8900	10,3600	0,6305	0,6163	0,6063	
1012214195-6944-cal-1.sam	22	56,1800	9,2800	0,6100	0,1800	1,4200	30,2800	59,8300	9,8900	0,6334	0,6199	0,6107	
1012127952-6944-cal-1.sam	23	47,8000	8,5500	1,1200	0,2800	1,4600	38,4700	51,8600	9,6700	0,4929	0,4804	0,4707	
1012214416-6950-Cal-1.sam	24	55,8200	8,8200	0,5800	0,2200	1,7100	31,0600	59,5400	9,4000	0,5992	0,5862	0,5763	
1012214355-6944-cal-1.sam(1)	25	56,6600	9,3300	0,6500	0,2100	1,6300	29,4700	60,5500	9,9800	0,6017	0,5882	0,5781	
1012125477-6949-Cal-1.sam	26	55,8200	11,2100	0,8200	0,2300	1,5800	28,1200	59,8400	12,0300	0,6298	0,6198	0,6114	
1012127617-6950-Cal-1.sam	27	59,1200	8,7400	0,4900	0,2100	1,9300	27,5400	63,2300	9,2300	0,6167	0,6038	0,5943	
1012125293-6944-cal-1.sam	28	52,0900	8,4100	1,0200	0,2100	1,4600	34,5300	56,0400	9,4300	0,5888	0,5783	0,5694	
101224850-6948-Cal-1.sam	29	57,0700	10,4300	0,7300	0,2100	1,4900	28,2100	60,6300	11,1600	0,5945	0,5819	0,5718	
1012215154-6949-Cal-1.sam	30	53,8700	10,5900	0,7600	0,2500	1,6200	30,5900	58,0600	11,3500	0,5669	0,5577	0,5498	
1012129390-6944-cal-1.sam	31	50,6600	8,5300	0,8800	0,3200	1,6700	35,6000	54,9800	9,4100	0,5157	0,5036	0,4934	
1012126658-6944-cal-1.sam	32	49,4200	8,0800	0,7400	0,3800	1,6200	37,7200	53,4700	8,8100	0,5358	0,5255	0,5174	
1012214874-6975-cal-1.sam	33	56,8500	7,4400	0,5800	0,2400	2,1700	30,7200	61,2600	8,0100	0,6086	0,5955	0,5854	
1012215215-6972-cal-1.sam	34	54,7700	8,7000	0,5400	0,2400	1,5700	32,1500	58,6100	9,2400	0,5336	0,5196	0,5094	
1012127730-6968-cal-1.sam	35	57,0800	7,7100	0,6400	0,2400	1,6500	30,2600	61,3900	8,3500	0,5620	0,5517	0,5433	
1012214690-6966-cal-1.sam	36	55,6000	8,5700	0,9700	0,3100	1,9200	30,4800	59,9800	9,5400	0,5897	0,5769	0,5676	

Name : Matriz para trabajo  
Size : 159X149  
Created : 27/04/2023 13:

Info Notes

## Matriz para tra...

# The construction of the matrix

- Data import
- Reduce (average)
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- **Insert Category variable**
- Define range for the purpose of grouping rows and columns
- Plot line





- Insert Category variable

Espectros Uva NIRCROP 2023 [Data Matrix uva.unsbj] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Espectros Uva NIRCROP Data Matrix

Data Matrix	Fecha	Variedad	pH	BRIX	Acidez (%)	908.1	914.294	920.489	926.683	932.877	939.072	945.266	951.4
	1	2	3	4	5	6	7	8	9	10	11	12	
19 07 Temp	1	19-07	Tempranillo	3,4200	18,1			8823	0,8723	0,8630	0,8559	0,8537	0,8555
19 07 Temp	2	19-07	Tempranillo	3,2200	15,3			8848	0,8686	0,8545	0,8430	0,8397	0,8435
19 07 Temp	3	19-07	Tempranillo	3,2500	13,7			8869	0,8694	0,8516	0,8378	0,8340	0,8376
19 07 Temp	4	19-07	Tempranillo	3,2300	15,9			9045	0,8884	0,8735	0,8627	0,8605	0,8622
19 07 Temp	5	19-07	Tempranillo	2,6500	7,7			7540	0,7323	0,7120	0,6974	0,6928	0,6958
19 07 Temp	6	19-07	Tempranillo	3,2000	17,1			9456	0,9307	0,9153	0,9053	0,9034	0,9079
19 07 Temp	7	19-07	Tempranillo	2,8000	14,4			8223	0,8051	0,7860	0,7721	0,7686	0,7699
19 07 Temp	8	19-07	Tempranillo	2,8800	17,6			8541	0,8374	0,8213	0,8094	0,8047	0,8083
19 07 Temp	9	19-07	Tempranillo	3,0400	14,4			8581	0,8418	0,8259	0,8123	0,8074	0,8095
19 07 Temp	10	19-07	Tempranillo	2,8700	11,3			7934	0,7716	0,7503	0,7355	0,7302	0,7335
19 07 Temp	11	19-07	Tempranillo	2,6200	13,0			8060	0,7865	0,7666	0,7524	0,7485	0,7505
19 07 Temp	12	19-07	Tempranillo	3,1400	12,4			8812	0,8640	0,8470	0,8325	0,8284	0,8303
19 07 Temp	13	19-07	Tempranillo	2,9500	16,6			8707	0,8521	0,8355	0,8231	0,8188	0,8235
19 07 Temp	14	19-07	Tempranillo	2,8100	14,4			7967	0,7773	0,7567	0,7412	0,7361	0,7362
19 07 Temp	15	19-07	Tempranillo	2,9600	18,1			8851	0,8677	0,8516	0,8379	0,8347	0,8372
19 07 Temp	16	19-07	Tempranillo	2,8800	15,9			9004	0,8856	0,8716	0,8601	0,8550	0,8546
19 07 Temp	17	19-07	Tempranillo	2,7900	12,7						0,8693	0,8650	0,8652
19 07 Temp	18	19-07	Tempranillo	2,8300	14,3						0,7486	0,7426	0,7449
19 07 Temp	19	19-07	Tempranillo	2,8600	16,3						0,8201	0,8139	0,8141
19 07 Temp	20	19-07	Tempranillo	2,5400	12,2			8301	0,8127	0,7961	0,7821	0,7757	0,7766
19 07 Temp	21	19-07	Tempranillo	3,7500	19,4			9051	0,8897	0,8742	0,8607	0,8565	0,8588
19 07 Temp	22	19-07	Tempranillo	3,1100	13,0			8732	0,8544	0,8384	0,8258	0,8206	0,8237
19 07 Temp	23	19-07	Tempranillo	3,0900	14,6			9485	0,9331	0,9171	0,9048	0,9014	0,9036
19 07 Temp	24	19-07	Tempranillo	3,2500	16,9			9098	0,8927	0,8776	0,8661	0,8616	0,8646
19 07 Temp	25	19-07	Tempranillo	3,0900	14,3000	0,4970	0,9410	0,9258	0,9120	0,8996	0,8882	0,8832	0,8832
19 07 Temp	26	19-07	Tempranillo	3,0900	15,6600	0,4600	0,9596	0,9456	0,9314	0,9169	0,9081	0,9032	0,9052
19 07 Temp	27	19-07	Tempranillo	2,9800	15,0500	0,4130	0,9251	0,9115	0,8978	0,8848	0,8748	0,8729	0,8736
19 07 Temp	28	19-07	Tempranillo	2,9500	17,7700	0,3820	0,9372	0,9210	0,9068	0,8931	0,8836	0,8798	0,8837
19 07 Temp	29	19-07	Tempranillo	2,8600	13,6900	0,6130	0,8702	0,8510	0,8314	0,8135	0,7994	0,7931	0,7927
19 07 Temp	30	19-07	Tempranillo	3,0300	17,2100	0,3550	0,9194	0,9033	0,8869	0,8721	0,8597	0,8556	0,8550

Context menu options:

- Create Column Range
- Undo
- Redo
- Cut
- Copy
- Copy with Headers
- Paste
- Insert Copied Cells
- Append Copied Cells
- Split Text/Category Variable
- Delete...
- Fill...
- Insert
  - Row(s)/Column(s)...
  - Category Variable...
- Append
- Change Data Type
- Make Header
- Plot

Info Notes

Name : Data Matrix  
Size : 780X135  
Created : 15/04/2024 12:...

Calibration Data Matrix

Value: 18,13 71

- Espectros Uva NIRCROP 2023 [Data Matrix uva.unsb] - The Unscrambler X

Value: 18,13 78



- Insert Category variable

Espectros Uva NIRCROP 2023 [Data Matrix uva.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Espectros Uva NIRCROP Data Matrix

		1	2	3	4	5	6	7	9
		Fecha	Variedad	pH	Acidez (%)	908.1	914.294	914.294	914.294
19 07 Temp	1	19-07	Tempranillo	3,2400	18,1300	0,3960	0,8915	0,8823	
19 07 Temp	2	19-07	Tempranillo	3,2200	15,3800	0,3470	0,9024	0,8848	
19 07 Temp	3	19-07	Tempranillo	3,2500	13,7500	0,6760	0,9053	0,8869	
19 07 Temp	4	19-07	Tempranillo	3,2300	15,9800	0,3810	0,9207	0,9045	
19 07 Temp	5	19-07	Macabeo	2,6500	7,7300	1,0490	0,7748	0,7540	
19 07 Temp	6	19-07	Tempranillo	3,2000	17,1900	0,3310	0,9609	0,9456	
19 07 Temp	7	19-07	Tempranillo	2,8000	14,4500	0,6160	0,8418	0,8223	
19 07 Temp	8	19-07	Tempranillo	2,8800	17,6600	0,7260	0,8708	0,8541	
19 07 Temp	9	19-07	Tempranillo	3,0400	14,4500	0,5610	0,8761	0,8581	
19 07 Temp	10	19-07	Tempranillo	2,8700	11,3700	0,4180	0,8154	0,7934	
19 07 Temp	11	19-07	Tempranillo	2,6200	13,0100	0,8600	0,8253	0,8060	
19 07 Temp	12	19-07	Tempranillo	3,1400	12,4700	0,5370	0,8998	0,8812	
19 07 Temp	13	19-07	Tempranillo	2,9500	16,6700	0,3980	0,8898	0,8707	
19 07 Temp	14	19-07	Tempranillo	2,8100	14,4800	0,5020	0,8168	0,7967	
19 07 Temp	15	19-07	Tempranillo	2,9600	18,1500	0,2800	0,9015	0,8851	
19 07 Temp	16	19-07	Tempranillo	2,8800	15,9300	0,2580	0,9169	0,9004	
19 07 Temp	17	19-07	Tempranillo	2,7900	12,7000	0,6090	0,9249	0,9089	
19 07 Temp	18	19-07	Tempranillo	2,8300	14,3800	1,2220	0,8227	0,8027	
19 07 Temp	19	19-07	Tempranillo	2,8600	16,3700	0,5870	0,8860	0,8684	
19 07 Temp	20	19-07	Tempranillo	2,5400	12,2400	0,7860	0,8472	0,8301	
19 07 Temp	21	19-07	Tempranillo	3,7500	19,4700	0,3970	0,9218	0,9051	
19 07 Temp	22	19-07	Tempranillo	3,1100	13,0500	0,5130	0,8924	0,8732	
19 07 Temp	23	19-07	Tempranillo	3,0900	14,6600	0,5340	0,9649	0,9485	
19 07 Temp	24	19-07	Tempranillo	3,2500	16,9700	0,2830	0,9285	0,9098	
19 07 Temp	25	19-07	Tempranillo	3,0900	14,3000	0,4970	0,9410	0,9258	
19 07 Temp	26	19-07	Tempranillo	3,0900	15,6600	0,4600	0,9596	0,9456	
19 07 Temp	27	19-07	Tempranillo	2,9800	15,0500	0,4130	0,9251	0,9115	
19 07 Temp	28	19-07	Tempranillo	2,9500	17,7700	0,3820	0,9372	0,9210	
19 07 Temp	29	19-07	Tempranillo	2,8600	13,6900	0,6130	0,8702	0,8510	
19 07 Temp	30	19-07	Tempranillo	3,0300	17,2100	0,3550	0,9194	0,9033	

Name : Data Matrix  
Size : 780X135  
Created : 15/04/2024 12:...

Info Notes

Calibration Data Matrix

Category Variable

Category Name: Variedad

Method

☒ Specify the level manually

☐ Specify levels to be based on a collection of row set

Category Name: Macabeo Add

Tempranillo  
Macabeo Remove

Up Down

OK Cancel

Right side/fill

# The construction of the matrix

- Data import
- Reduce (average)
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- **Define range for the purpose of grouping rows and columns**
- Plot line



- Define range for the purpose of grouping rows and columns

Espectros Uva NIRCROP 2023 [Data Matrix uva.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Espectros Uva NIRCROP Data Matrix

Transform  
Plot  
Export  
Range  
Duplicate Matrix...  
Rename  
Delete  
Spectra  
Save Matrix

Define Range...  
Copy Range  
Paste Range

Fecha	Variedad	pH	BRIX	Acidez (%)	908.1	914.294	920.489	926.683
19-07	Tempranillo	3,4200	18,1300	0,3960	0,8915	0,8823	0,8723	0,86
19-07	Tempranillo	3,2200	15,3800	0,3470	0,9024	0,8848	0,8686	0,85
19-07	Tempranillo	3,2500	13,7500	0,6760	0,9053	0,8869	0,8694	0,85
19-07	Tempranillo	3,2300	15,9800	0,3810	0,9207	0,9045	0,8884	0,87
19-07	Tempranillo	2,6500	7,7300	1,0490	0,7748	0,7540	0,7323	0,71
19-07	Tempranillo	3,2000	17,1900	0,3310	0,9609	0,9456	0,9307	0,91
19-07	Tempranillo	2,8000	14,4500	0,6160	0,8418	0,8223	0,8051	0,78
19-07	Tempranillo	2,8800	17,6600	0,7260	0,8708	0,8541	0,8374	0,82
19-07	Tempranillo	3,0400	14,4500	0,5610	0,8761			
19-07	Tempranillo	2,8700	11,3700	0,4180	0,8158			
19-07	Tempranillo	2,6200	13,0100	0,8600	0,8668			
19-07	Tempranillo	3,1400	12,4700	0,5370	0,8668			
19-07	Tempranillo	2,9500	16,6700	0,3980	0,8823			
19-07	Tempranillo	2,8100	14,4800	0,5020	0,8168			
19-07	Tempranillo	2,9600	18,1500	0,2800	0,9015			
19-07	Tempranillo	2,8800	15,9300	0,2580	0,9169			
19-07	Tempranillo	2,7900	12,7000	0,6090	0,9249			
19-07	Tempranillo	2,8300	14,3800	1,2220	0,8227			
19-07	Tempranillo	2,8600	16,3700	0,5870	0,8860			
19-07	Tempranillo	2,5400	12,2400	0,7860	0,8472			
19-07	Tempranillo	3,7500	19,4700	0,3970	0,9218			
19-07	Tempranillo	3,1100	13,0500	0,5130	0,8924			
19-07	Tempranillo	3,0900	14,6600	0,5340	0,9649			
19-07	Tempranillo	3,2500	16,9700	0,2830	0,9285			
19-07	Tempranillo	3,0900	14,3000	0,4970	0,9410			
19-07	Tempranillo	3,0900	15,6600	0,4600	0,9596			
19-07	Tempranillo	2,9800	15,0500	0,4130	0,9251			
19-07	Tempranillo	2,9500	17,7700	0,3820	0,9372			
19-07	Tempranillo	2,8600	13,6900	0,6130	0,8702			
19-07	Tempranillo	3,0300	17,2100	0,3550	0,9194			

Name: Data Matrix  
Size: 780X135  
Created: 15/04/2024 12:...

Info Notes

Calibration Data Matrix

Define Range

Row ranges

Macabeo	61-120,181
Tempranillo	1-60,121-1
Calibration	1-8,11-12,
External validation	9-10,13-14

Column ranges

Sello de tiempo	135
Espectros	6-130
Número de serie d	131
Temperatura	132
Notas	133
Brix	4
pH(1)	3
pH(2)	3

Keep out

Rows

Columns 1-2,131,133-135

Special intervals

Rowset

Update Create

Espectros 6-130

Update Create

Data Matrix	Variedad	pH	BRIX	Acidez (%)	908.1	914.294	920.489	926.683
19 07 Temp	1	Tempranillo	3,4200	18,1300	0,3960	0,8915	0,8823	0,8723
19 07 Temp	2	Tempranillo	3,2200	15,3800	0,3470	0,9024	0,8848	0,8686
19 07 Temp	3	Tempranillo	3,2500	13,7500	0,6760	0,9053	0,8869	0,8694
19 07 Temp	4	Tempranillo	3,2300	15,9800	0,3810	0,9207	0,9045	0,8884
19 07 Temp	5	Tempranillo	2,6500	7,7300	1,0490	0,7748	0,7540	0,7323
19 07 Temp	6	Tempranillo	3,2000	17,1900	0,3310	0,9609	0,9456	0,9307
19 07 Temp	7	Tempranillo	2,8000	14,4500	0,6160	0,8418	0,8223	0,8051
19 07 Temp	8	Tempranillo	2,8800	17,6600	0,7260	0,8708	0,8541	0,8374

Size: 780X135

Selected columns: 125

OK Cancel



- Define range for the purpose of grouping rows and columns

Espectros Uva NIRCROP 2023 [Data Matrix uva.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Data Matrix

Transform  
Plot  
Export  
Range  
Duplicate Matrix...  
Rename  
Delete  
Spectra  
Save Matrix

Define Range...

Copy Range  
Paste Range

Fecha	Variedad	pH	BRIX	Acidez (%)	908.1	914.294
19-07	Tempranillo	3,4200	18,1300	0,3960	0,8915	0,8823
19-07	Tempranillo	3,2200	15,3800	0,3470	0,9024	0,8848
19-07	Tempranillo	3,2500	13,7500	0,6760	0,9053	0,8869
19-07	Tempranillo	3,2300	15,9800	0,3810	0,9207	0,9045
19-07	Tempranillo	2,6500	7,7300	1,0490	0,7748	0,7540
19-07	Tempranillo	3,2000	17,1900	0,3310	0,9609	0,9456
19-07	Tempranillo	2,8000	14,4500	0,6160	0,8418	0,8223
19-07	Tempranillo	2,8800	15,9300	0,2580	0,9169	0,8823
19-07	Tempranillo	2,7900	12,7000	0,6090	0,9249	0,8848
19-07	Tempranillo	2,8300	14,3800	1,2220	0,8227	0,8869
19-07	Tempranillo	2,8600	16,3700	0,5870	0,8860	0,9045
19-07	Tempranillo	2,5400	12,2400	0,7860	0,8472	0,7540
19-07	Tempranillo	3,7500	19,4700	0,3970	0,9218	0,9456
19-07	Tempranillo	3,1100	13,0500	0,5130	0,8924	0,8223
19-07	Tempranillo	3,0900	14,6600	0,5340	0,9649	0,8541
19-07	Tempranillo	3,2500	16,9700	0,2830	0,9285	0,8374
19-07	Tempranillo	3,0900	14,3000	0,4970	0,9410	0,8051
19-07	Tempranillo	3,0900	15,6600	0,4600	0,9596	0,8051
19-07	Tempranillo	2,9800	15,0500	0,4130	0,9251	0,8051
19-07	Tempranillo	2,9500	17,7700	0,3820	0,9372	0,8051
19-07	Tempranillo	2,8600	13,6900	0,6130	0,8702	0,8051
19-07	Tempranillo	3,0300	17,2100	0,3550	0,9194	0,8051

Name: Data Matrix  
Size: 780X135  
Created: 15/04/2024 12:...

Info Notes

Calibration Data Matrix

## Essential Ranges:

- Row range: calibration and validation
- Column range: set of spectral variables

## Define Range

Row ranges

Macabeo	61-120,181
Tempranillo	1-60,121-1
Calibration	1-8,11-12,
External validation	9-10,13-14

Column ranges

Sello de tiempo	135
Espectros	6-130
Número de serie d	131
Temperatura	132
Notas	133
Brix	4
pH(1)	3

Keep out

Rows

Columns 1-2,131,133-135

Special intervals

Create

Size: 780X135

Selected columns: 125

- Define range for the purpose of grouping rows and columns

### Calibration and validation sets

Define Range

Row ranges  
validación con mu 175-196

Column ranges  
All spectra 15-139  
Oleico 10

Calibration 1,4,7,13,16,19,22,25

Columnset

Update Create

Keep out  
Rows  
Columns

**Special intervals**

Interval  
Every 3 Sample  
Starting from 1

Random  
Sample 0

Category

OK Cancel

Size: 196X146 Selected rows: 66

Abs	ID muestra	Set	C12:0	C14:0	C16:0	C1
	1	2	3	4	5	
1012125590-	1	1,0000	Calibración	0,0600	1,1900	20,5300
1012214775-	2	2,0000	Calibración	0,0700	1,1000	18,4100
1012214171-	3	3,0000	Calibración	0,0600	1,1900	18,3600
1012125910-	4	4,0000	Calibración	0,0700	1,3100	19,9900
1012128256-	5	5,0000	Calibración	0,0700	1,3700	21,1600
1012214317-	6	6,0000	Calibración	0,0600	1,1400	19,2900
1012127259-	7	7,0000	Calibración	0,0700	1,3600	19,5100
1012128416-	8	8,0000	Calibración	0,0600	1,0800	18,1100

- Define range for the purpose of grouping rows and columns

### Calibration and validation sets

Define Range

Row ranges

Macabeo	61-100,181
Tempranillo	60,121-1
Calibration	1-8,11-12,
External validation	9-10,13-14

Column ranges

Sello de tiempo	135
Espectros	6-130
Número de serie d	131
Temperatura	132
Notas	133
Brix	4
pH(1)	3
pH(2)	3

Keep out

Rows

Columns 1-2,131,133-135

Special intervals

Interval

Every 0 Sample

Starting from 1

Random

Sample 0

Category

Calibration 1-8,11-12,15-22,25-26,2

Columnset

Update Create

Data Matrix		Fecha	Variedad	pH	BRIX	Acidez (%)	90
		1	2	3	4	5	
19 07 Temp	1	19-07	Tempranillo	3,4200	18,1300	0,3960	
19 07 Temp	2	19-07	Tempranillo	3,2200	15,3800	0,3470	
19 07 Temp	3	19-07	Tempranillo	3,2500	13,7500	0,6760	
19 07 Temp	4	19-07	Tempranillo	3,2300	15,9800	0,3810	
19 07 Temp	5	19-07	Tempranillo	2,6500	7,7300	1,0490	
19 07 Temp	6	19-07	Tempranillo	3,2000	17,1900	0,3310	
19 07 Temp	7	19-07	Tempranillo	2,8000	14,4500	0,6160	
19 07 Temp	8	19-07	Tempranillo	2,8800	17,6600	0,7260	

Size: 780X135

Selected rows: 550

It is also possible to define calibration and validation sets by means of a category variable



# The construction of the matrix

- Data import
- Reduce (average)
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)
- **Plot line**

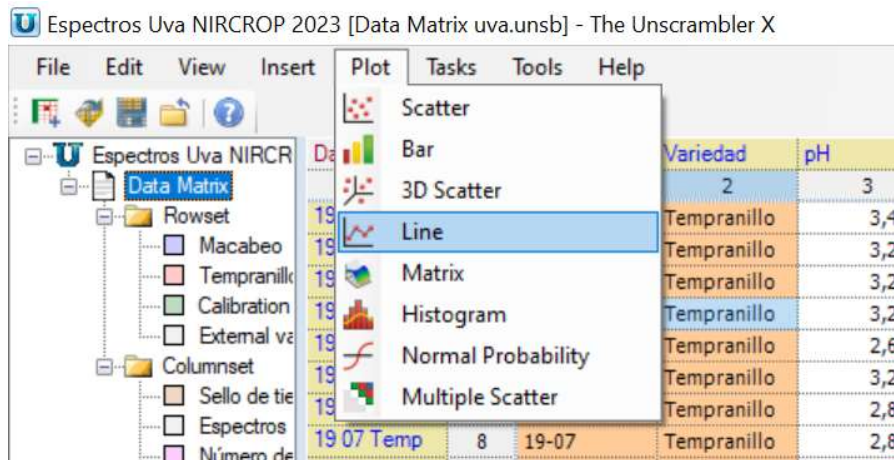


- Plot line

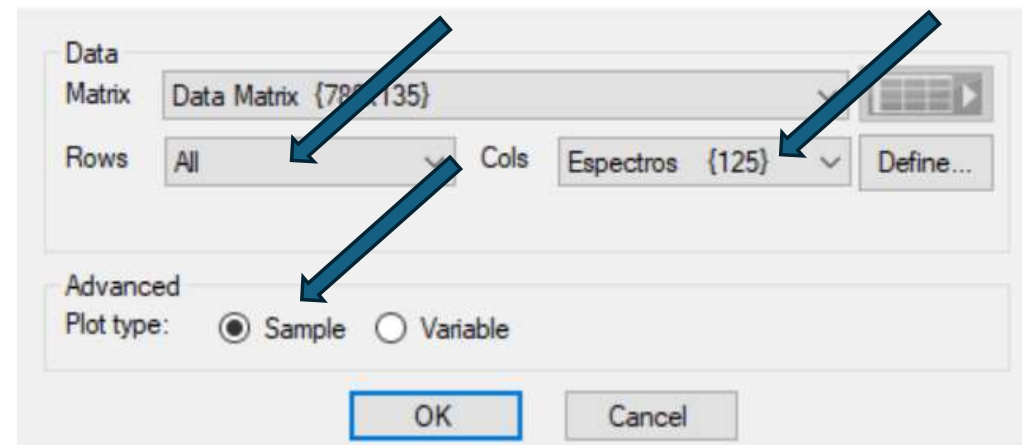


**In order to visualize the spectra and detect possible anomalies**

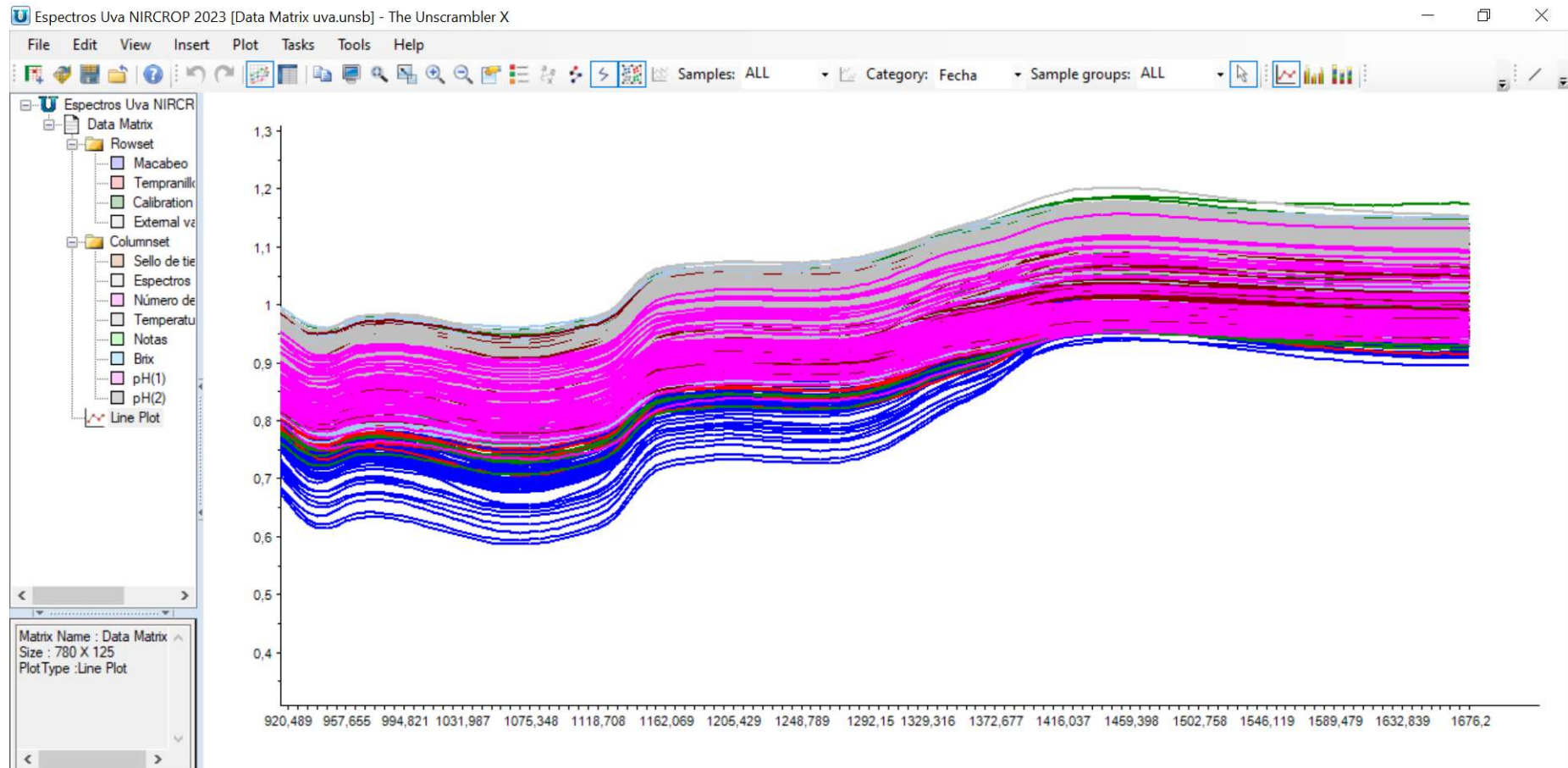
## Plot/line

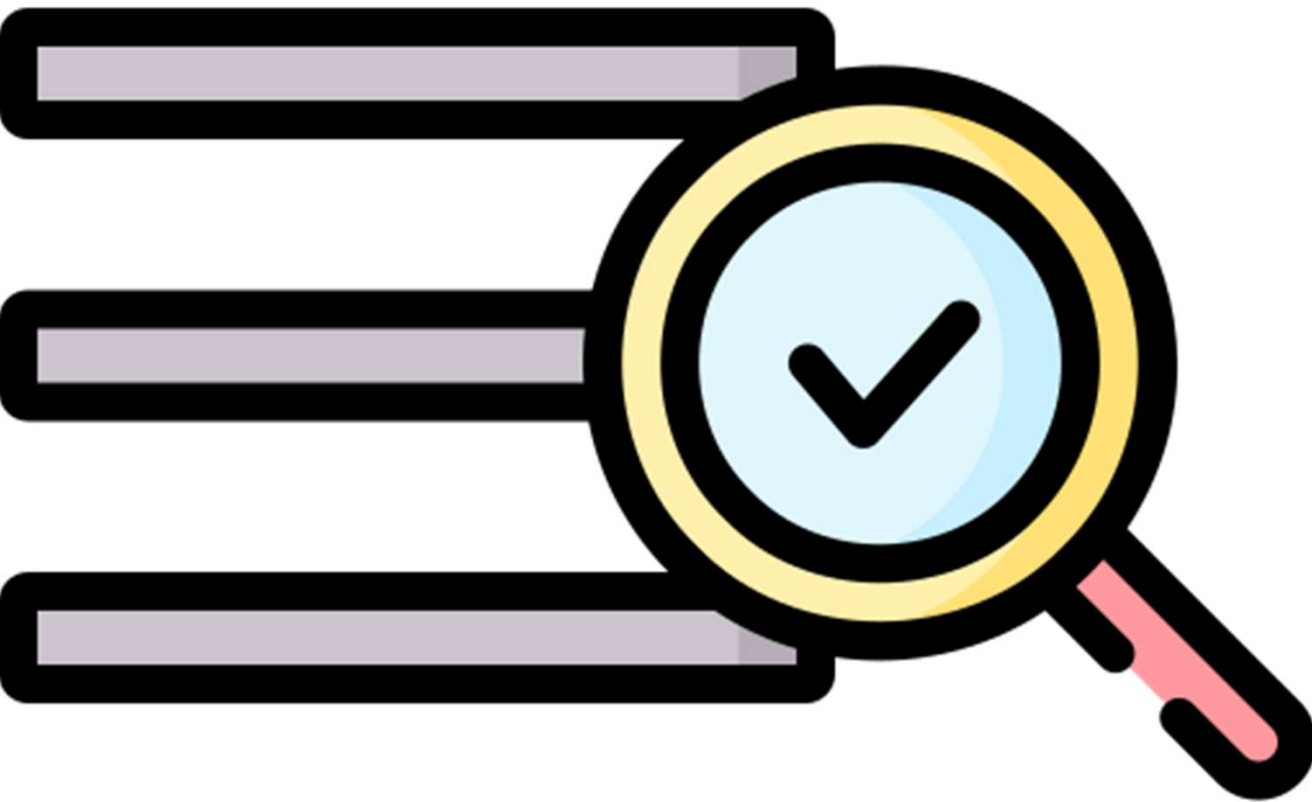


## Line Plot - Scope



- Plot line





## Index

- What is NIRS technology?
- The Unscrambler
  - The construction of the matrix
  - **Quantitative models**
  - Qualitative models
  - Pre-treatments



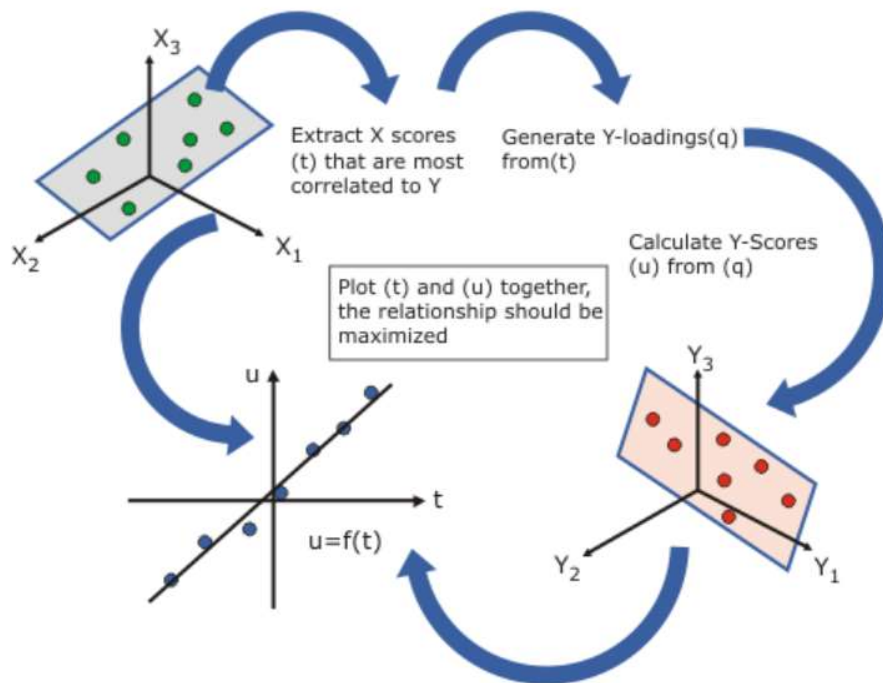
## Quantitative models

- **Construction of the model by means of Partial least squares regression (PLSR)**
- Methods available for validation
- Plots
- Statistics for assessing a quantitative model
- Predict
- Outliers



- Construction of the model by means of Partial least squares regression (PLSR)

**Partial Least Squares regression** models both the X- and Y-matrices simultaneously **to find the latent variables in X that will best predict the latent variables in Y**. These PLSR components are similar to principal components; however, they are referred to as *factors*. (It finds a linear regression model by projecting the predicted variables and the observable variables to a new space).



- Construction of the model by means of Partial least squares regression (PLSR)

**Partial Least Squares regression is a multivariate analysis, i.e. it uses multiple variables - spectral variables - to estimate a given physicochemical parameter.**

**Why cannot a univariate approach be performed in NIRS?**

- ☐ **Overlapping absorption bands of different sample constituents**
- ☐ **Large amount of data**
- ☐ **High redundant information (collinearity)**



**Causes low selectivity of NIR data**

- Construction of the model by means of Partial least squares regression (PLSR)

## Task/Analyze/Partial least square regression

**Data matrix:** the same for predictors and responses

The screenshot shows the 'TOTAL SPECTRA LONCHADOS' data matrix in The Unscrambler X. The data matrix table has columns for 'Fecha', 'Variedad', 'pH', 'BRIX', 'Acidez (%)', and numerical values. The 'Model Inputs' dialog box is open, showing 'Data Matrix {780x135}' for both Predictors (X) and Responses (Y). The 'Rows' are set to 'Calibration {550}' and 'Cols' are 'Espectros {125}' for X and 'Brix {1}' for Y. The 'Maximum components' is set to 10. The 'Identify outliers' and 'Mean center data' options are checked. The 'Finish' button is highlighted.

**Row range: calibration or another specific row range**

**Number of maximum components: max. 10% of number of samples**

Builds a model based on Partial least squares

- Construction of the model by means of Partial least squares regression (PLSR)

Partial Least Squares Regression

Setup

- Model Inputs
- X Weights
- Y Weights
- Validation
- Algorithm
- Outlier limits

Validation

☐ Leverage correction

☒ Cross validation

☐ Prediction diagnostics for CV segments

☒ Uncertainty test

Number of components in uncertainty test

☒ Use optimal number

☐ Use set number

1

Data

X

Rows

Cols

Define...

Data

Y

Rows

Cols

Define...

Cancel Back Next Finish

Uncertainty test/Use optimal number

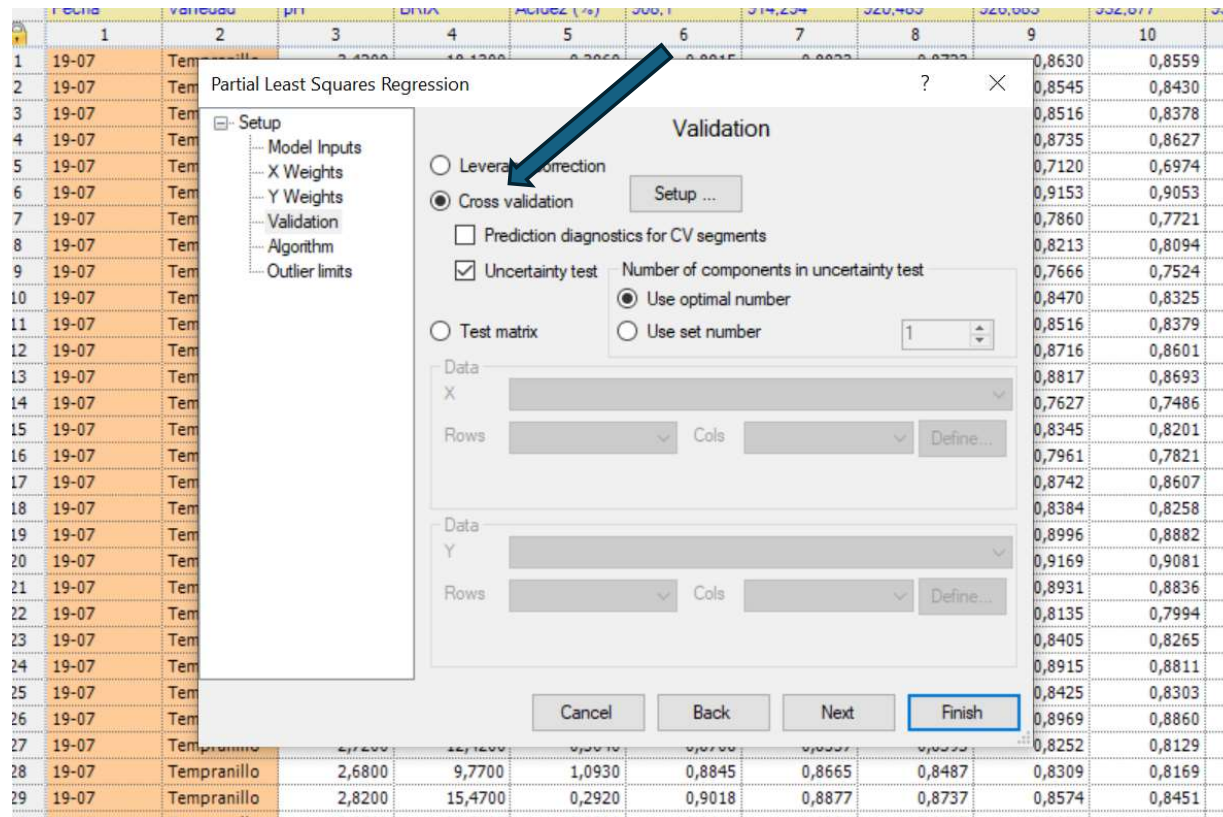


## Quantitative models

- Construction of the model by means of Partial least squares regression (PLSR)
- **Methods available for validation**
- Plots
- Statistics for assessing a quantitative model
- Predict
- Outliers



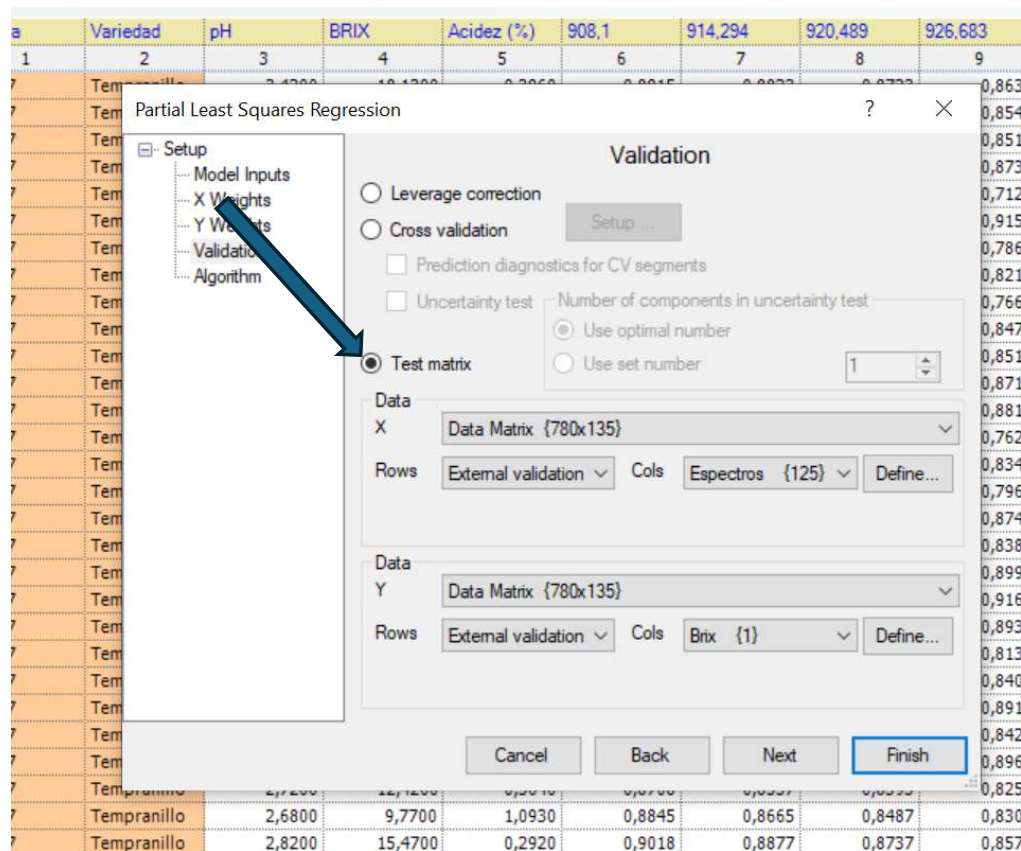
- Methods available for validation



## Cross-Validation

This method is used when either **there are not enough samples available to make a separate test set**, or for simulating the effects of different validation test cases, e.g. systematically leaving samples out vs. randomly leaving samples out, etc.

- Methods available for validation



## Test matrix

This is also known as *Test Set Validation* and uses independent samples that have not taken part in the calibration for validation. Both X- and Y-matrices need to be defined in this case. **This is the preferred method for validation and should be aimed for.**

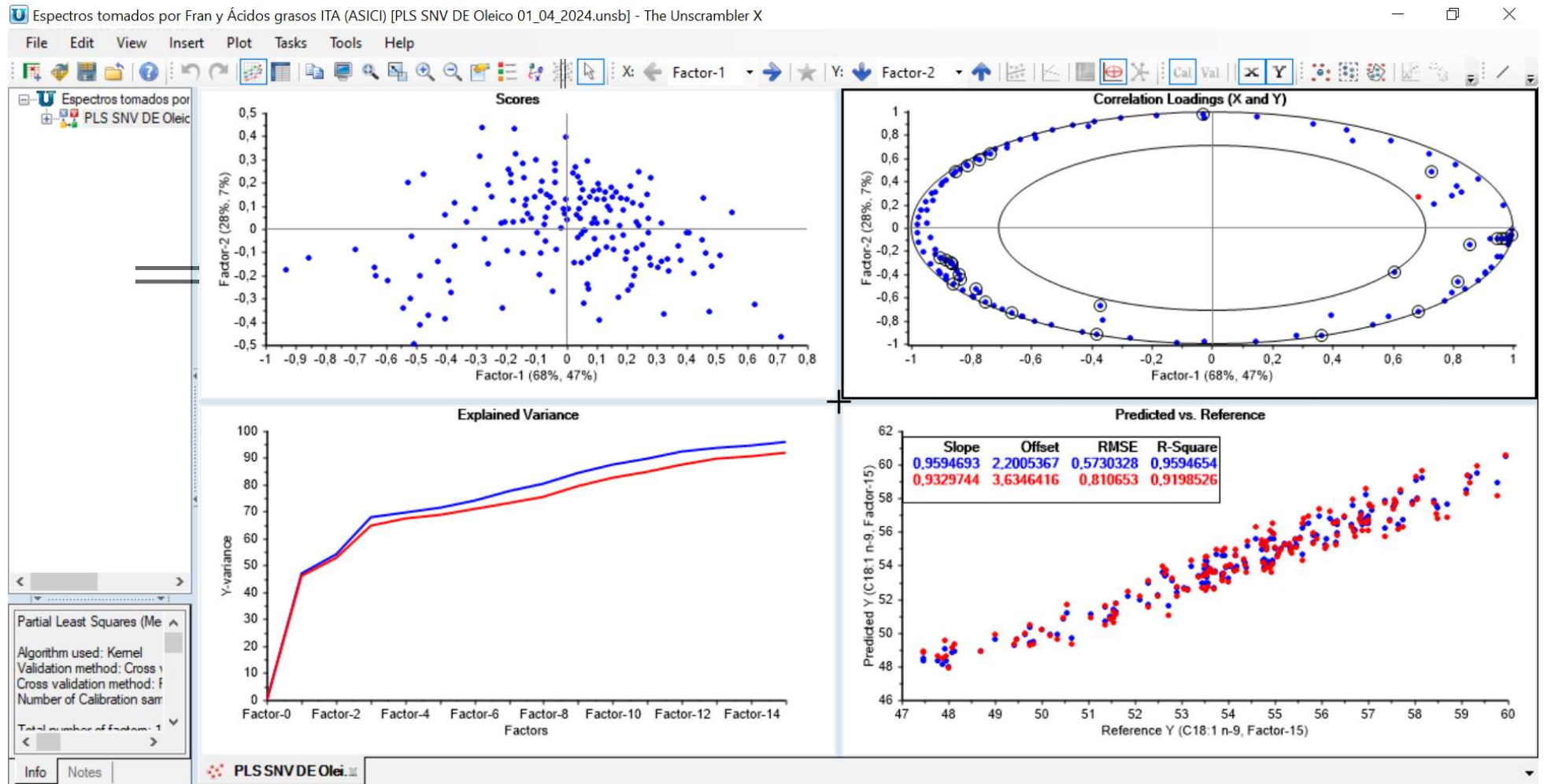


## Quantitative models

- Construction of the model by means of Partial least squares regression (PLSR)
- Methods available for validation
- **Plots**
- Statistics for assessing a quantitative model
- Predict
- Outliers

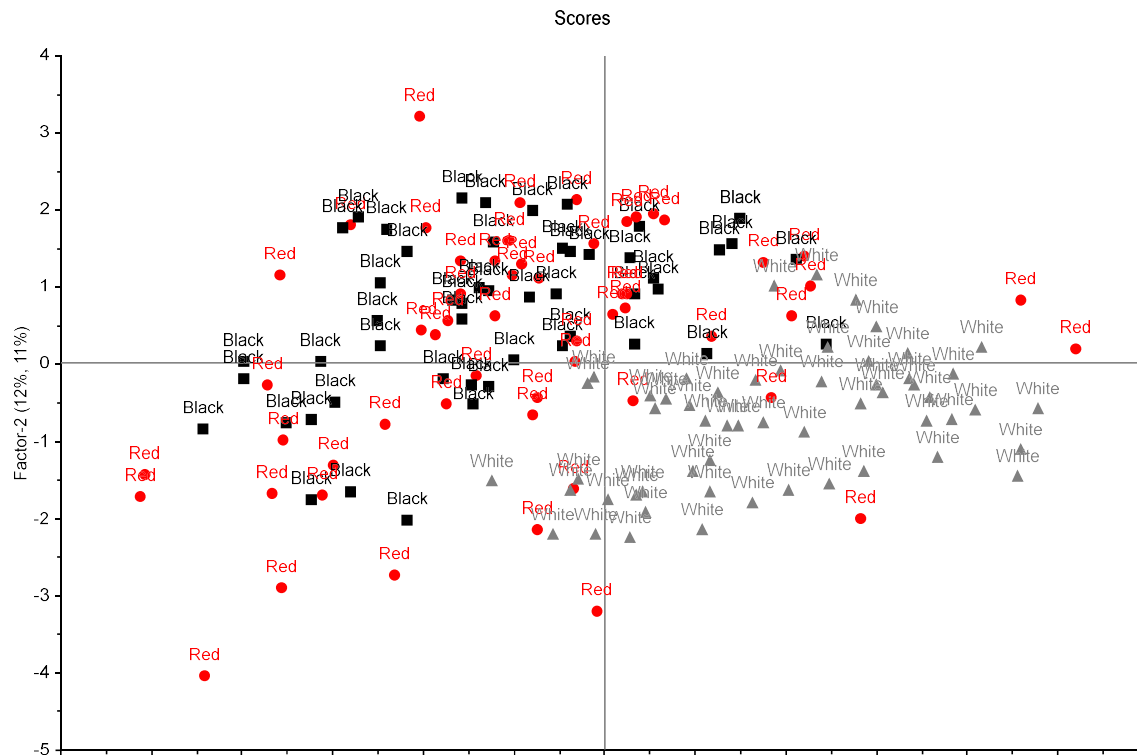
- Plots

## Regression Overview



- Plots

## Scores plot



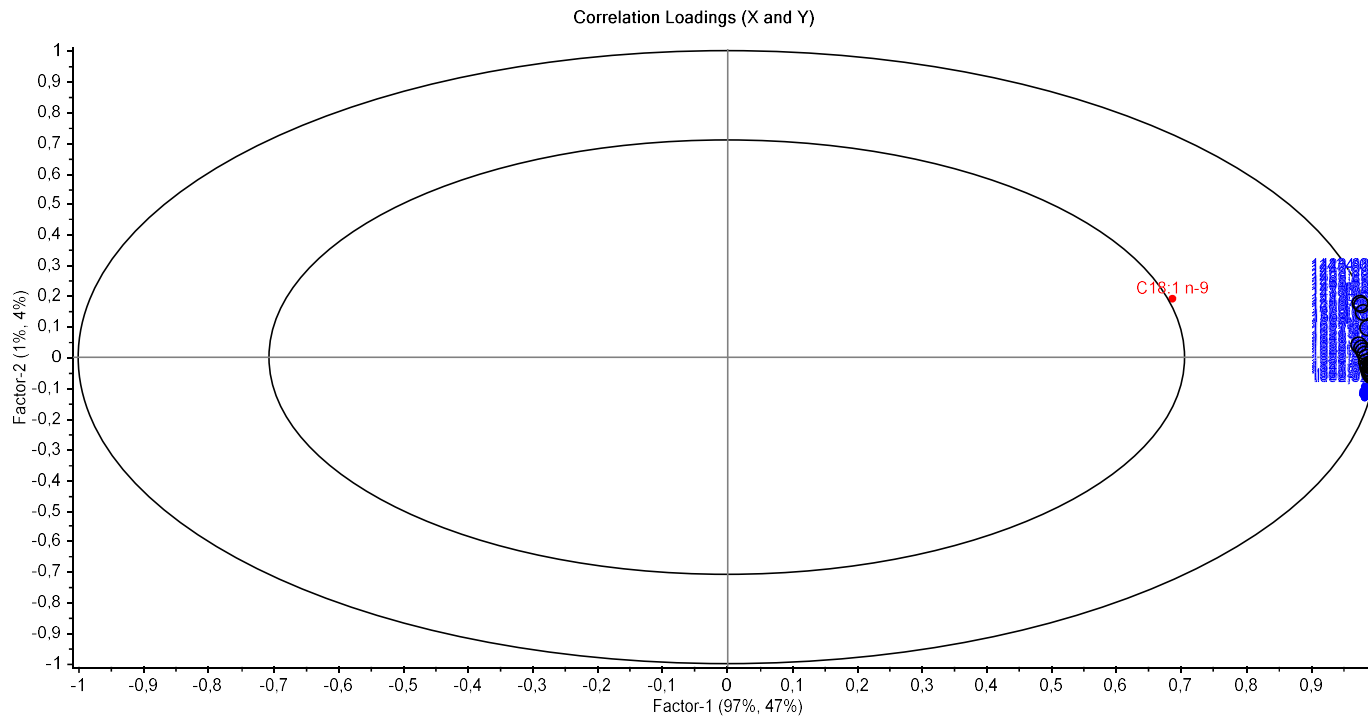
This is a two-dimensional scatter plot (or map) of scores for two specified factors (latent variables or PCs) from PLS regression.

The plot gives information about patterns in the samples. The scores plot for (factor 1, factor 2) is especially useful, since these two components summarize more variation in the data than any other pair of components. **The closer the samples are in the scores plot, the more similar they are with respect to the two components concerned.** Conversely, samples far away from each other are different from each other, so the plot can be used to interpret differences and similarities among samples. It allows to study sample distribution or even detect an outlier.



- Plots

## Correlation loadings

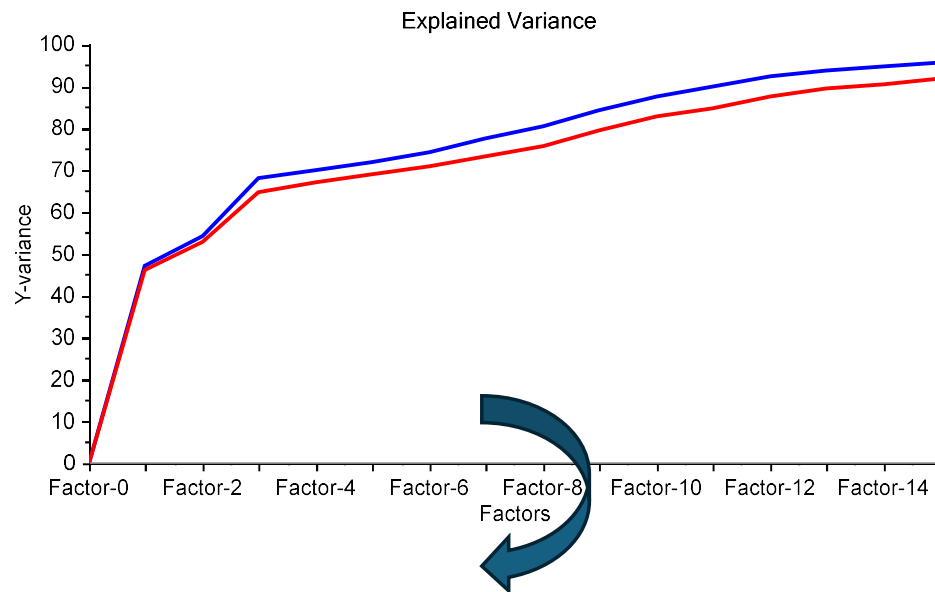


A 2-D scatter plot of X- and Y-loadings for two specified components (factors) from PLS is a **good way to detect important variables and relationships between variables**. The two ellipses help to check how much variance is taken into account. The outer ellipse indicates 100% explained variance. The inner ellipse indicates 50% of explained variance.

- Predictors (X) projected in roughly the same direction from the center as a response, are positively linked to that response.
- Predictors projected in the opposite direction have a negative link.
- **Predictors projected close to the center, are not well represented in that model and cannot be interpreted.**

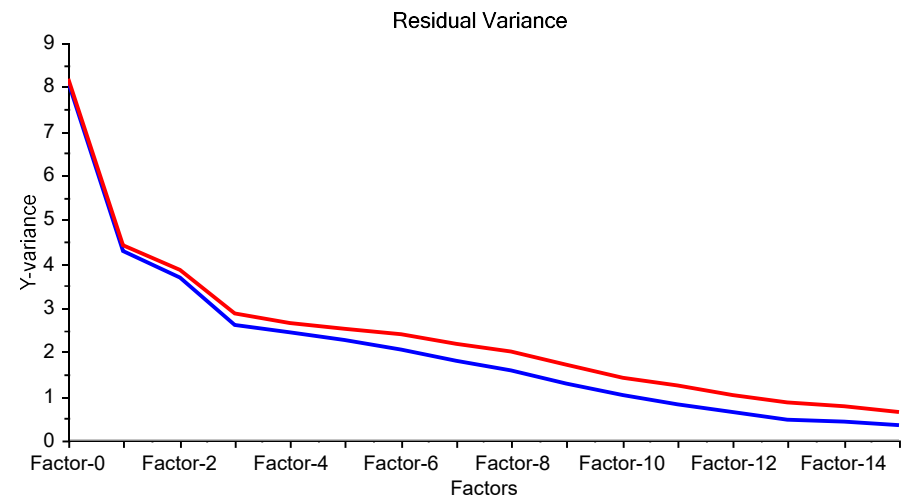
- Plots

### Explained/Residual variance variance



This plot illustrates how much of the variation in the response is described by each different component. **It is the percentage of the original variance in the data that is taken into account by the model**

### Residual variance

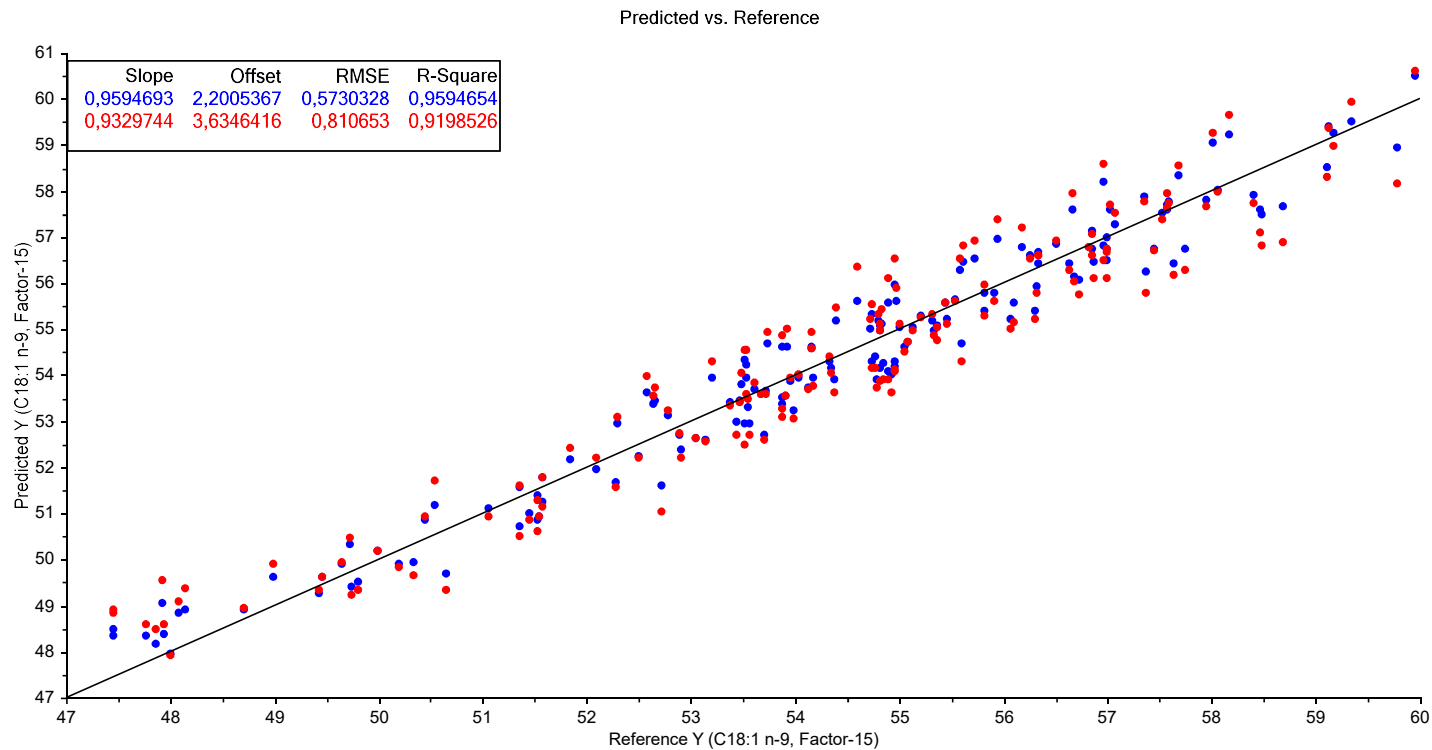


### Important!!!

Compare the two variances (calibration and validation): **if they differ significantly, there is good reason to question whether either the calibration data or the test data are truly representative!!**

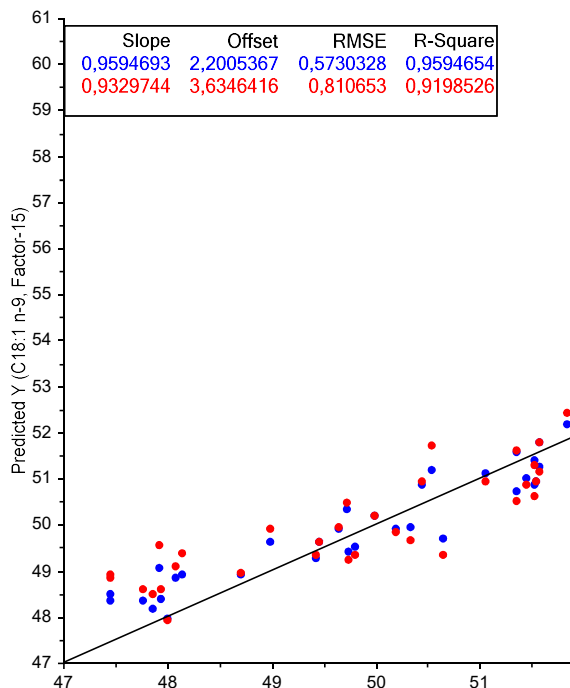
- Plots

## Predicted vs. Reference



The predicted Y-value from the model is plotted against the measured Y-value. This is a good way to check the quality of the regression model. **If the model gives a good fit, the plot will show points close to a straight line through the origin and with slope equal to 1**

- Plots



The predicted Y-value from the regression line is used to check the quality of the regression. The closer the predicted Y-value is to the actual Y-value, the closer to a straight line through the data points.

To determine the quality of the fit, the following statistics are available,

### Slope

The closer the slope is to **1**, the data are better modelled.

### Offset

This is the intercept of the line with the Y-axis when the X-axis is set to zero (**Note:** It is not a necessity that this value is zero!)

### RMSE

The first one (in blue) is the Calibration error. **This is a measure of the dispersion of the calibration samples about the regression line.** The second one (in red) is the expected Prediction error, depending on the validation method used. Both are expressed in the same unit as the response variable Y.

### R-squared

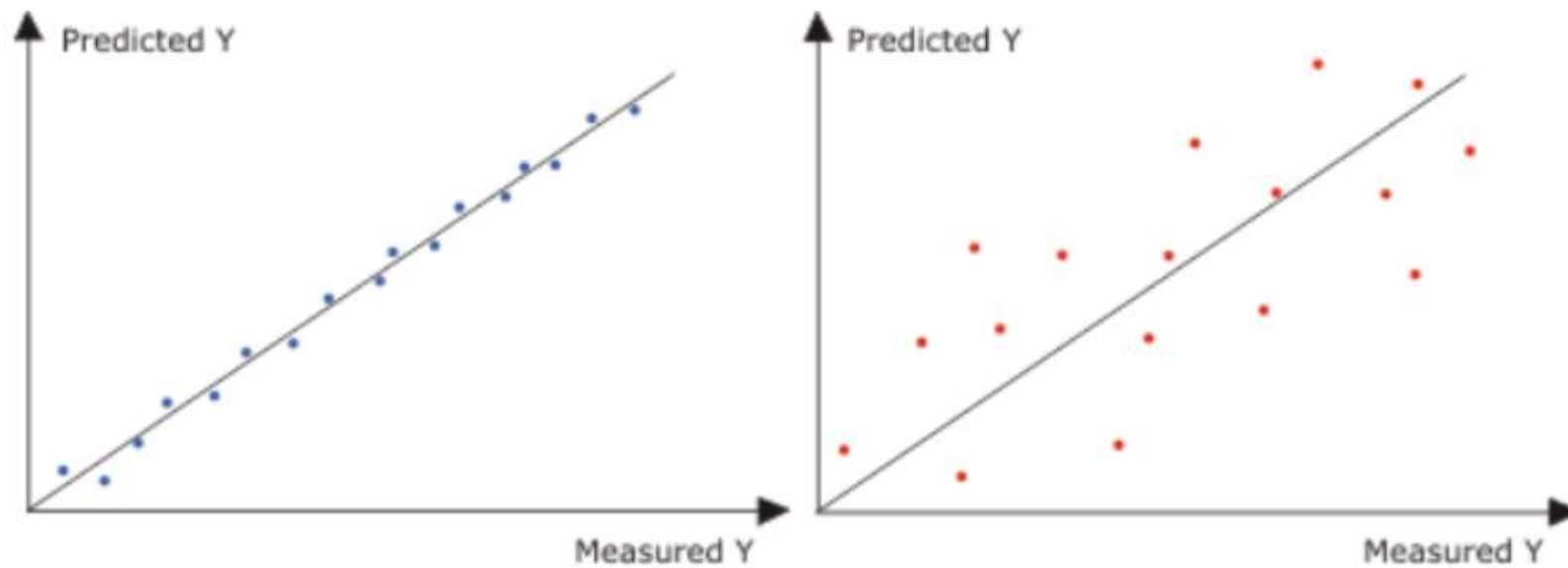
**It tells how good a fit can be expected for future predictions for a defined number of components.** The first one (in blue) is the calibration, the second one (in red) for the validation set.



- Plots

### Predicted vs. Reference

*Predicted vs. Reference shows how well the model fits*

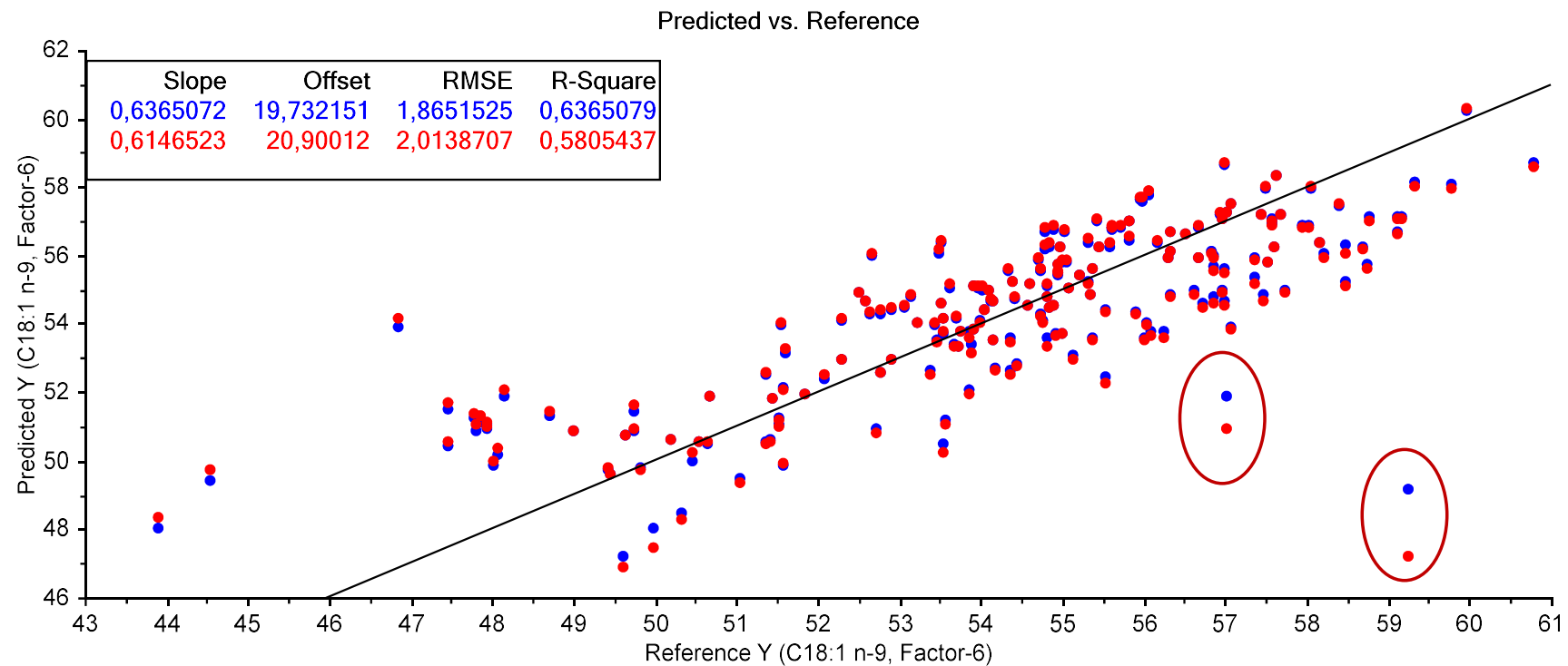


Left: Good fit. Right: Poor fit.

- Plots

## Predicted vs. Reference

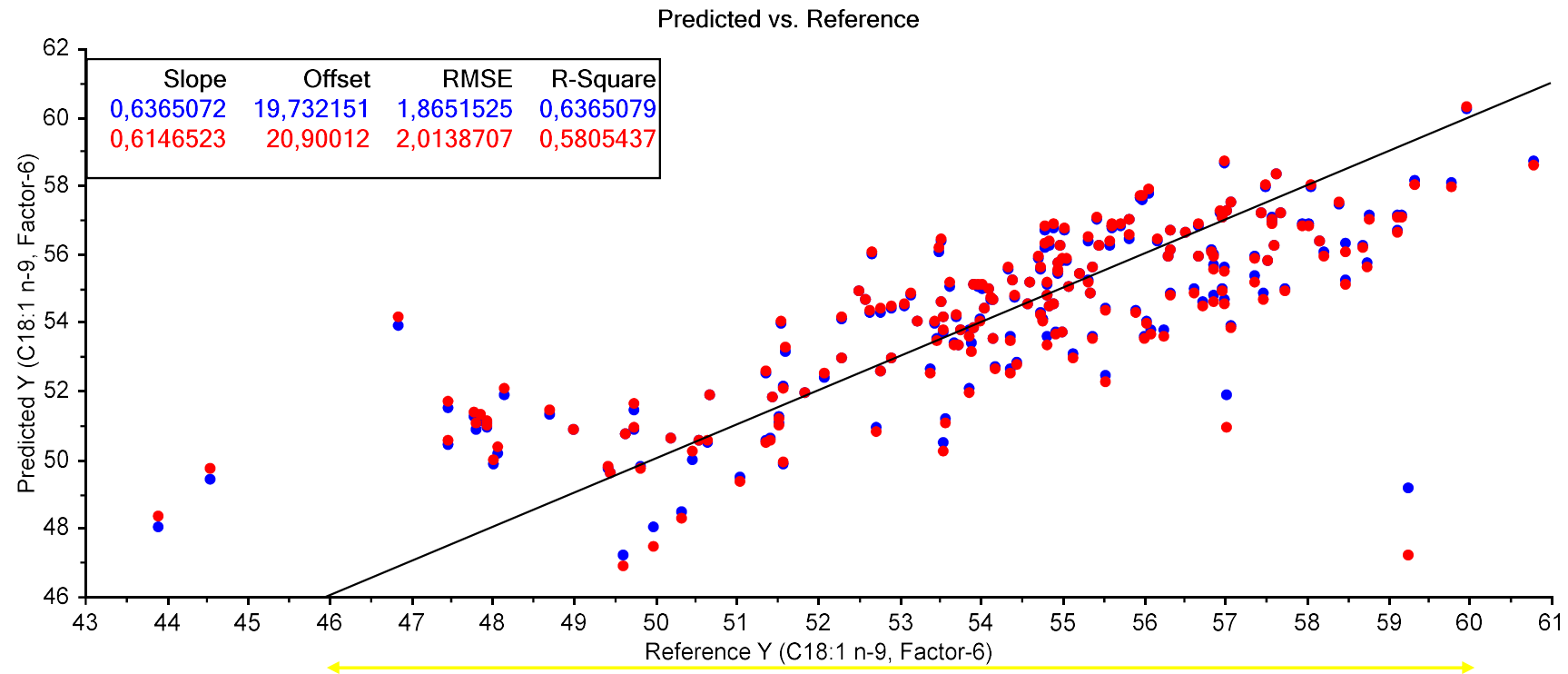
If there are large differences between the calibration and validation results (in cross-validation), the model cannot be trusted.



- Plots

## Predicted vs. Reference

**To achieve a good (robust, reliable) quantitative model, it is necessary to have a wide range of the variable to be predicted.**





## Quantitative models

- Construction of the model by means of Partial least squares regression (PLSR)
- Methods available for validation
- Plots
- **Statistics for assessing a quantitative model**
- Predict
- Outliers



- Statistics for assessing a quantitative model

### **Coeficiente de determinación en validación cruzada (1-VR)**

- ☐  $1-VR \geq 0,90$ . Excellent calibration
- ☐  $0,89 > 1-VR > 0,70$  Good calibration
- ☐  $0,69 > 1-VR > 0,50$  The model could be used to discriminate between low, medium and high values.

### **Error en validación cruzada**

Should be as little as possible

### **Residual Prediction Deviation (RPD)**

**$RPD = SD/RMSE$**

Routine analysis  $RPD \geq 3$

### **Range Error Ratio (RER)**

**$RER = Range/RMSE$**

- ☐  $4 > RER > 8$  The model could be used to discriminate between low, medium and high values
- ☐  $8 > RER > 12$  It is possible to predict quantitatively
- ☐  $RER > 12$  Excellent prediction



## Quantitative models

- Construction of the model by means of Partial least squares regression (PLSR)
- Methods available for validation
- Plots
- Statistics for assessing a quantitative model
- **Predict**
- Outliers

- Predict

**Use an existing regression model to predict response values for new samples.**

Tasks/Predict/Regression

The screenshot shows the 'Predict Using Regression Model' dialog box. It has a title bar with a question mark and a close button. The main area contains several sections: 'Select model' with a dropdown menu showing 'PLS SNV E Oleico 01\_04\_2024'; 'Components' with a dropdown menu showing '1'; a group of radio buttons for 'Short Prediction' (unselected) and 'Full Prediction' (selected); checkboxes for 'Inlier limit' (unchecked), 'Sample Inlier dist' (unchecked), and 'Identify Outliers' (checked); a 'Data' section with 'Matrix' set to 'Abs {196x146}' and 'Rows' set to 'Rowset Predict {1}'; and another 'Data' section with 'Matrix' set to 'Abs {196x146}' and 'Rows' set to 'Oleico {1}'. There are buttons for 'Pretreatment', 'Bias and Slope', 'Outlier Limits', 'Define...', 'OK', and 'Cancel'. Three blue arrows point to specific elements: one to the 'Full Prediction' radio button, one to the 'All spectra {125}' dropdown in the first 'Data' section, and one to the 'Oleico {1}' dropdown in the second 'Data' section.

☐ **Full Prediction** uses a projection on the latent space in the calculation. (It will provide comprehensive results such as plots and additional matrices for increased data interpretation and outlier diagnostics).

☐ **Short Prediction** uses only the extracted Regression (Beta) coefficients. There are no plots associated with this type of prediction.

- Predict

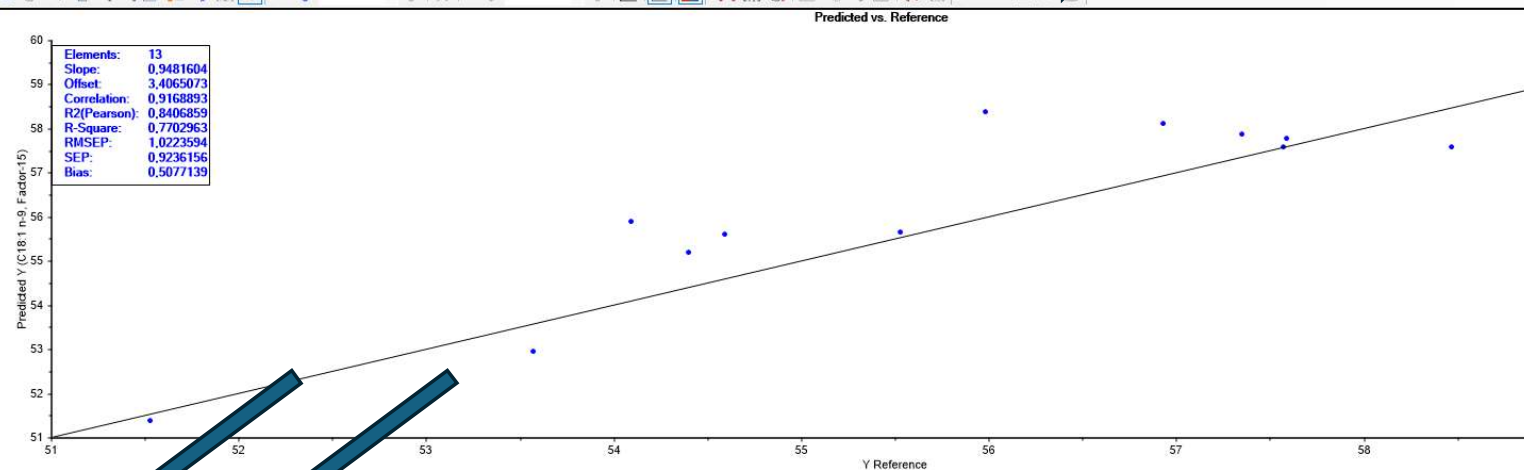
Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Matriz con espectros ampliación valores altos oleico para TRABAJAR.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

X: Factor-15 Y: C18:1 n-9

Espectros tomados por Fran y Ácidos grasos ITA (ASICI)

- Abs
  - Rowset
    - validación con muestras am
    - Rowset Predict
  - Columnset
    - All spectra
    - Oleico
  - PLS Oleico Abs
  - SNV\_Detrend
  - PLS SNV DE Oleico 01\_04\_2024
  - MSC\_Abs
  - Abs\_MSC
  - PLS MSC Oleico
  - SNV\_Detrend\_SG 1.4.4.1
  - PLS SNV DE SG 1.4.4.1 Oleico
  - Predict
  - Predict(1)
    - Raw data
    - Outputs
    - Sample Outliers
    - Plots
  - Predicted with Deviation
  - Predicted vs. Reference
  - Scores and Loadings
  - Residuals and Influence
  - Variable Contributions



Predicted vs. Reference

Predicted Y (C1)	Predicted	Deviation	Reference
1012125590-69	51,3945	0,6169	51,5300
1012214775-69	58,3832	0,8812	55,9800
1012214171-69	57,5959	0,7804	57,5700
1012125910-69	55,8975	0,8115	54,0900
1012128256-69	55,6468	0,6190	55,5300
1012214317-69	57,8851	0,6357	57,3500
1012127259-69	55,5967	0,6742	54,5900
1012128416-69	57,5953	0,8357	58,4700
1012214133-69	58,1092	1,0241	56,9300
1012214379-69	55,1877	0,6994	54,4000
1012214713-69	57,7921	0,9127	57,5900
1012215710-69	52,9601	0,6349	53,5700
1012213778-69	59,4963	0,8093	59,3400



## Quantitative models

- Construction of the model by means of Partial least squares regression (PLSR)
- Methods available for validation
- Plots
- Predict
- **Outliers**



- Outliers

An outlier is an **object which deviates from the other objects** in a model and may not **belong to the same population as the majority and therefore can disturb the model.**

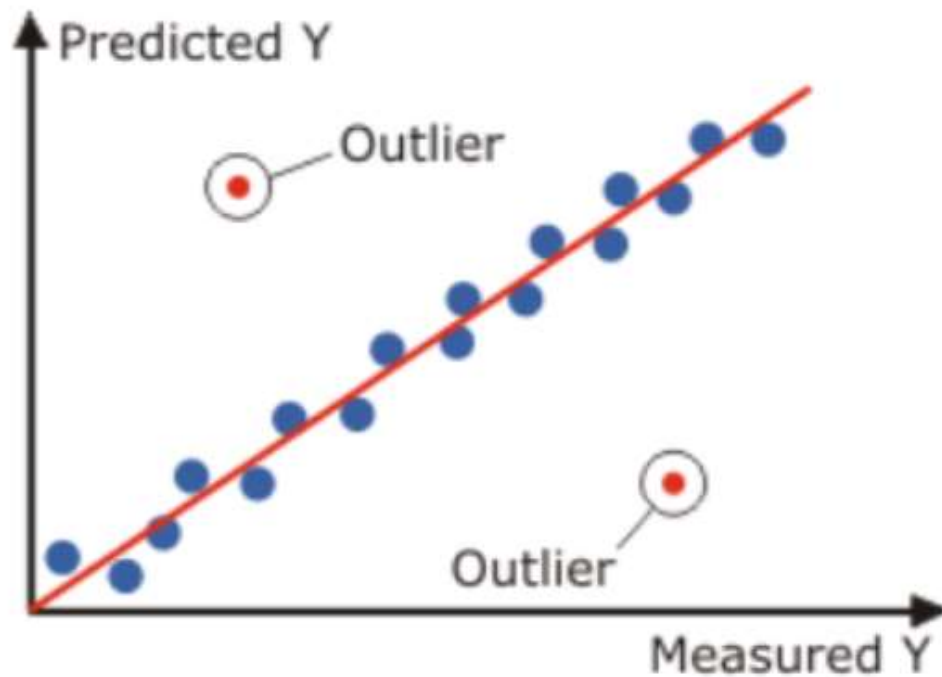
The cause of outliers could be one or more of the following:

- ❖ Measurement error
- ❖ Wrong labeling
- ❖ Deviating products / processes
- ❖ Noise
- ❖ **Extreme / interesting sample**

For projection methods like PLSR, outliers can be detected using scores plots, residuals, leverages and influence plots

- Outliers. **How to detect them?**

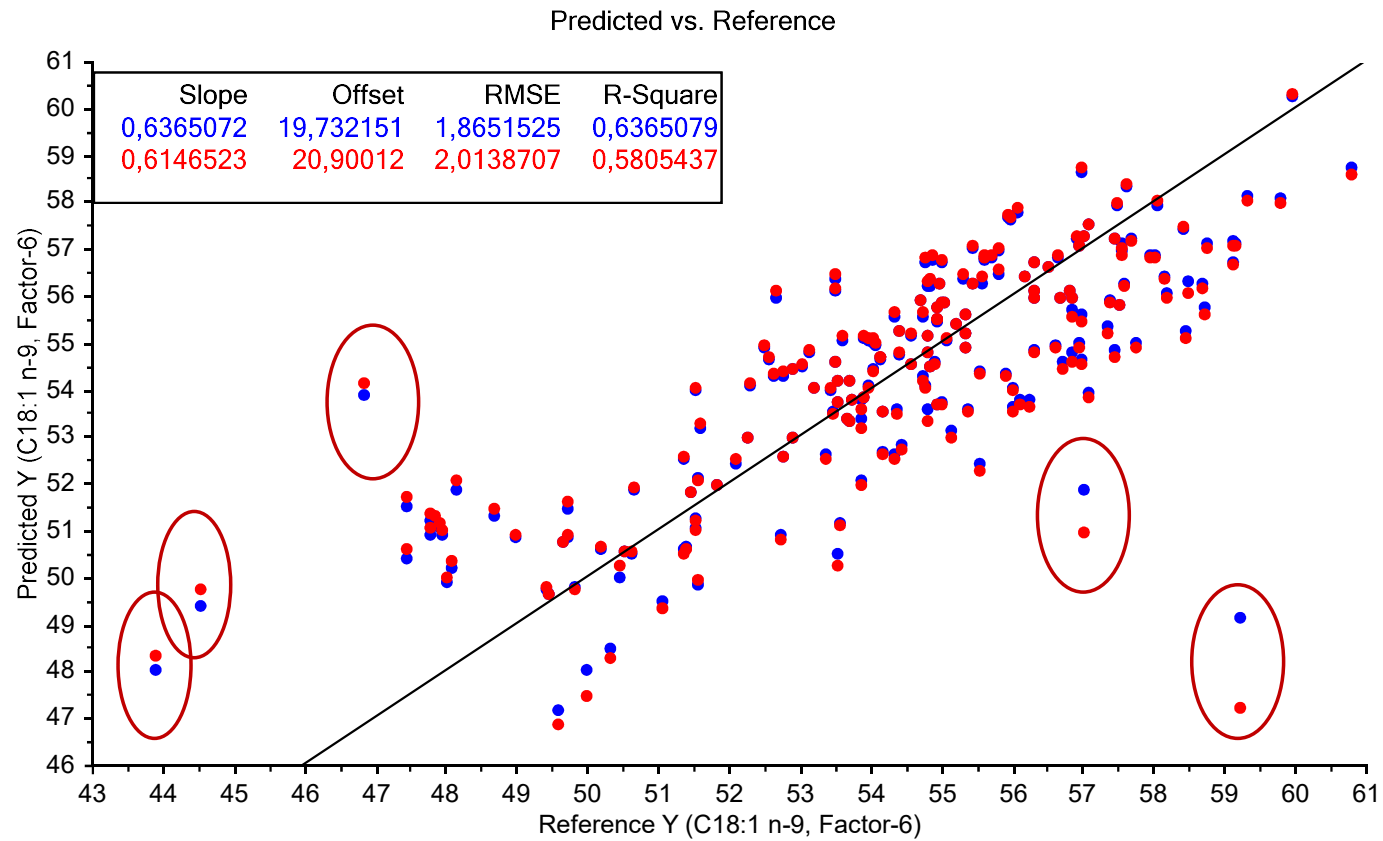
### Predicted vs. Reference



One may also see cases where the majority of the samples lie close to the line while a few of them are further away. This may indicate **good fit of the model to the majority of the data, but with a few outliers present**

- Outliers. **How to detect them?**

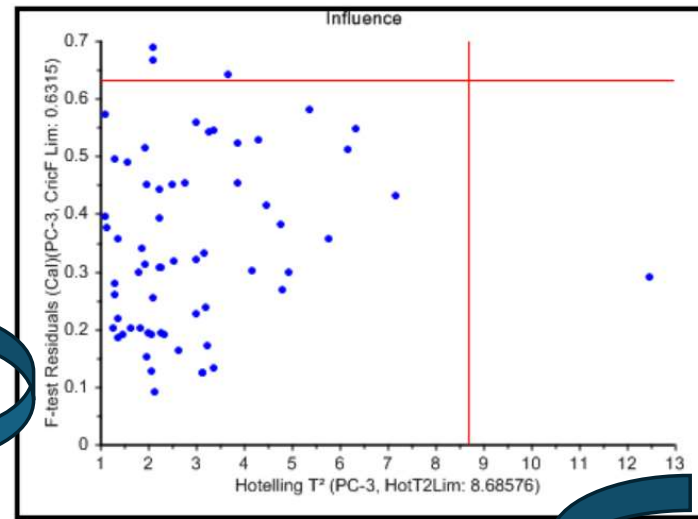
## Predicted vs. Reference



- Outliers. **How to detect them?**

**Influence plot with Hotelling's  $T^2$  statistic** ————— Representa dos tipos de outliers

*Influence plot with Hotelling's  $T^2$  on the abscissa and F-residuals on the ordinate*



El eje de ordenadas describe la distancia de la muestra al modelo.

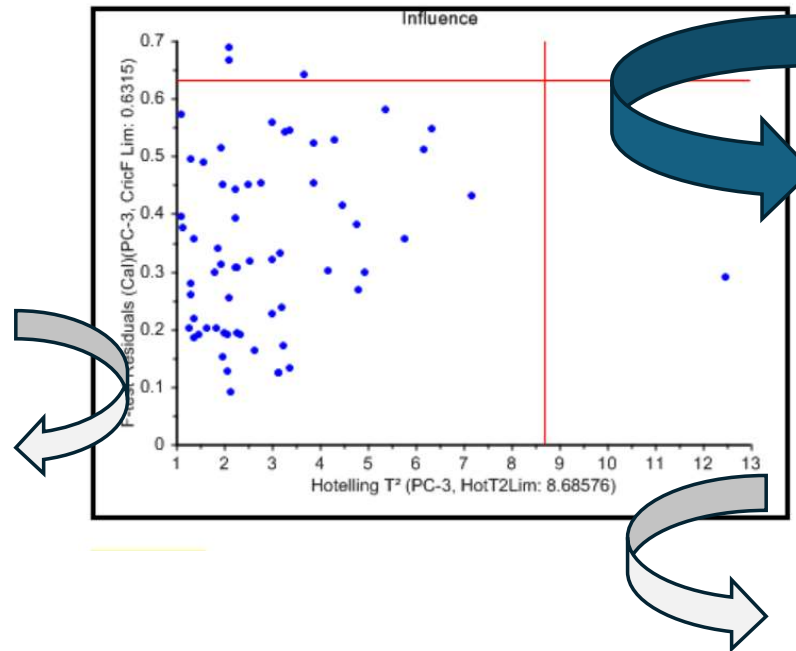
Las muestras con una varianza residual elevada están mal descritas por el modelo.

Las muestras con un valor elevado en el eje de abscisa (a la derecha) **están bien descritas por el modelo**. Sin embargo, las puntuaciones de las muestras pueden tener valores muy altos o bajos para algunos componentes en comparación con el resto de las muestras. Estas muestras son peligrosas en la fase de calibración porque influyen en el modelo. Una muestra suficientemente extrema puede abarcar por sí sola un componente entero, en cuyo caso el modelo dejará de ser fiable. **Si la varianza descrita por la muestra es importante pero única, hay que intentar obtener más muestras del mismo tipo para estabilizar el modelo. De lo contrario, la muestra debe descartarse como valor atípico.**

- Outliers. **How to detect them?**

**Influence plot with Hotelling's  $T^2$  statistic** ————— Representa dos tipos de outliers

*Influence plot with Hotelling's  $T^2$  on the abscissa and F-residuals on the ordinate*



**Obsérvese que una muestra con una varianza residual y un valor de Hotelling elevados es el valor atípico más peligroso. No sólo está mal descrita por el modelo, sino que además es influyente.**

componentes en comparación con el resto de las muestras. Estas muestras son peligrosas en la fase de calibración porque influyen en el modelo. Una muestra suficientemente extrema puede abarcar por sí sola un componente entero, en cuyo caso el modelo dejará de ser fiable. Si la varianza descrita por la muestra es importante pero única, hay que intentar obtener más muestras del mismo tipo para estabilizar el modelo. De lo contrario, la muestra debe descartarse como valor atípico.

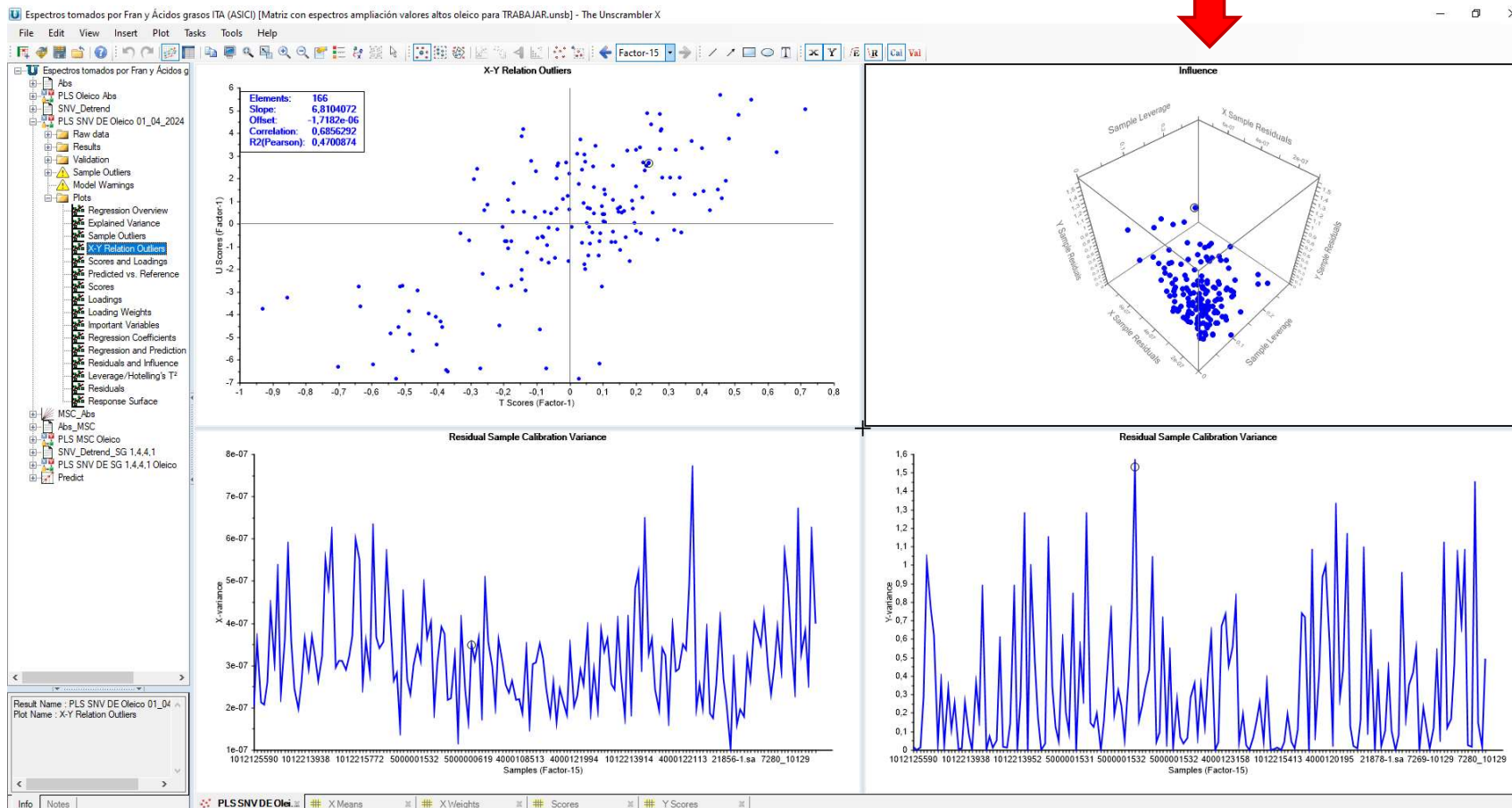
El eje de ordenadas describe la distancia de la muestra al modelo.

Las muestras con una varianza residual elevada están mal descritas por el modelo.



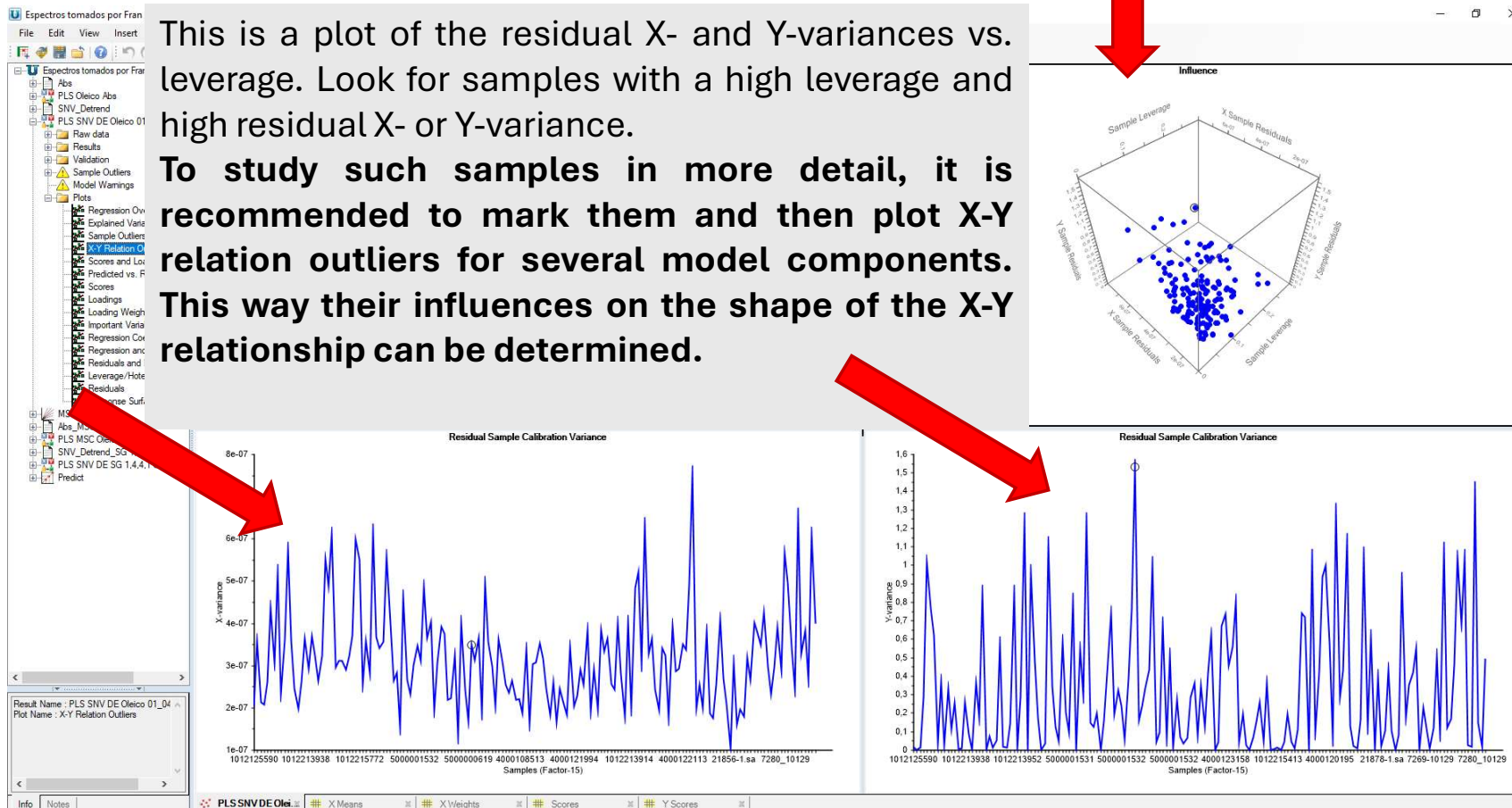
- Outliers. **How to detect them?**

## Influence plot



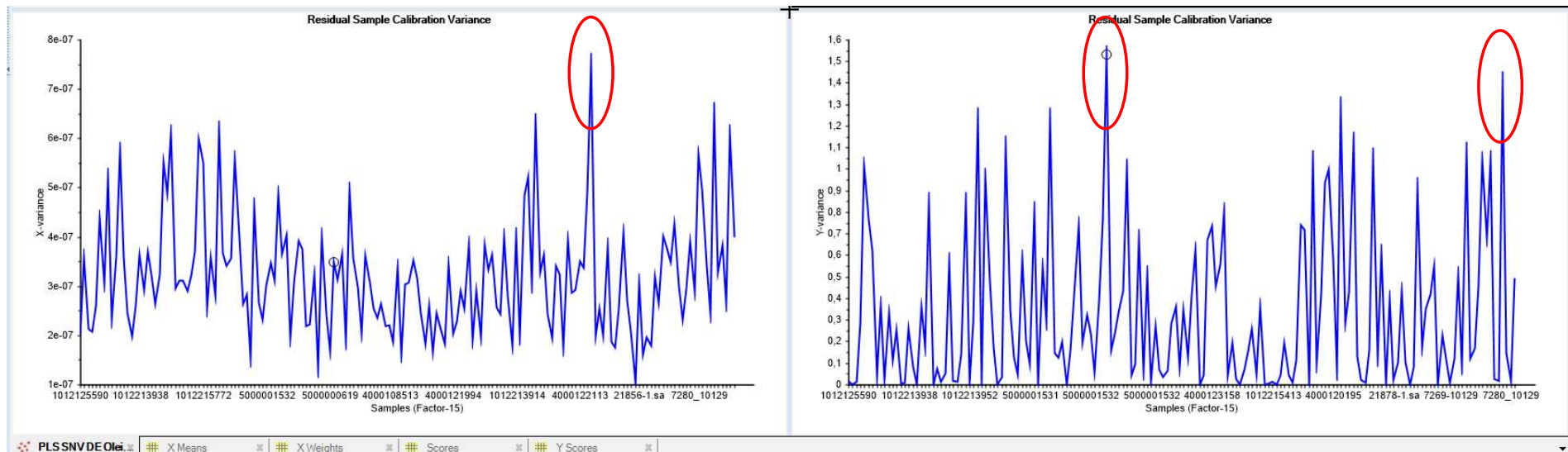
- Outliers. **How to detect them?**

## Influence plot



- Outliers. **How to detect them?**

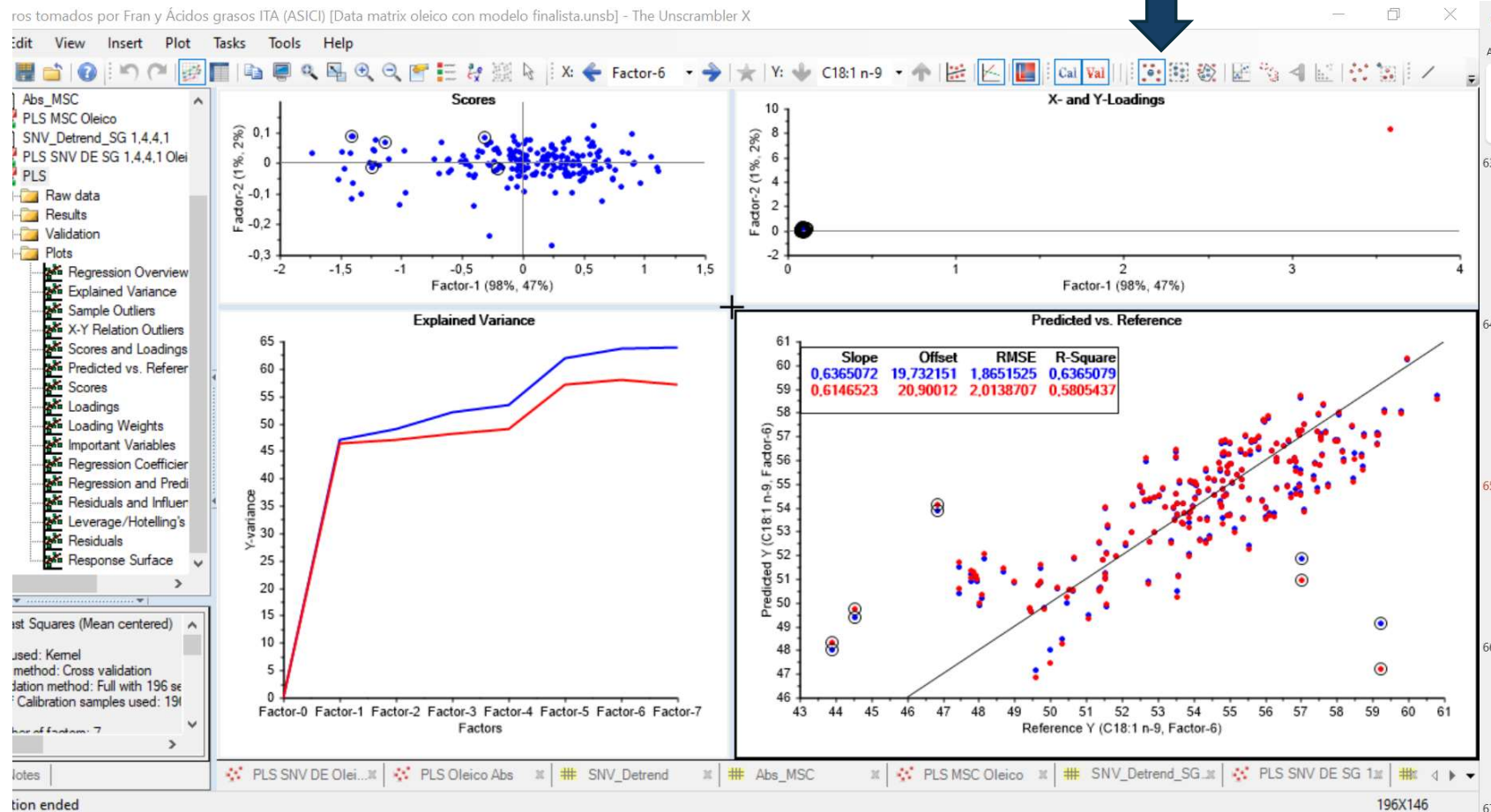
## Residual sample variance



High residuals indicate an outlier. Incorporating more components can sometimes model outliers; avoid doing so since it will reduce the prediction ability of the model.

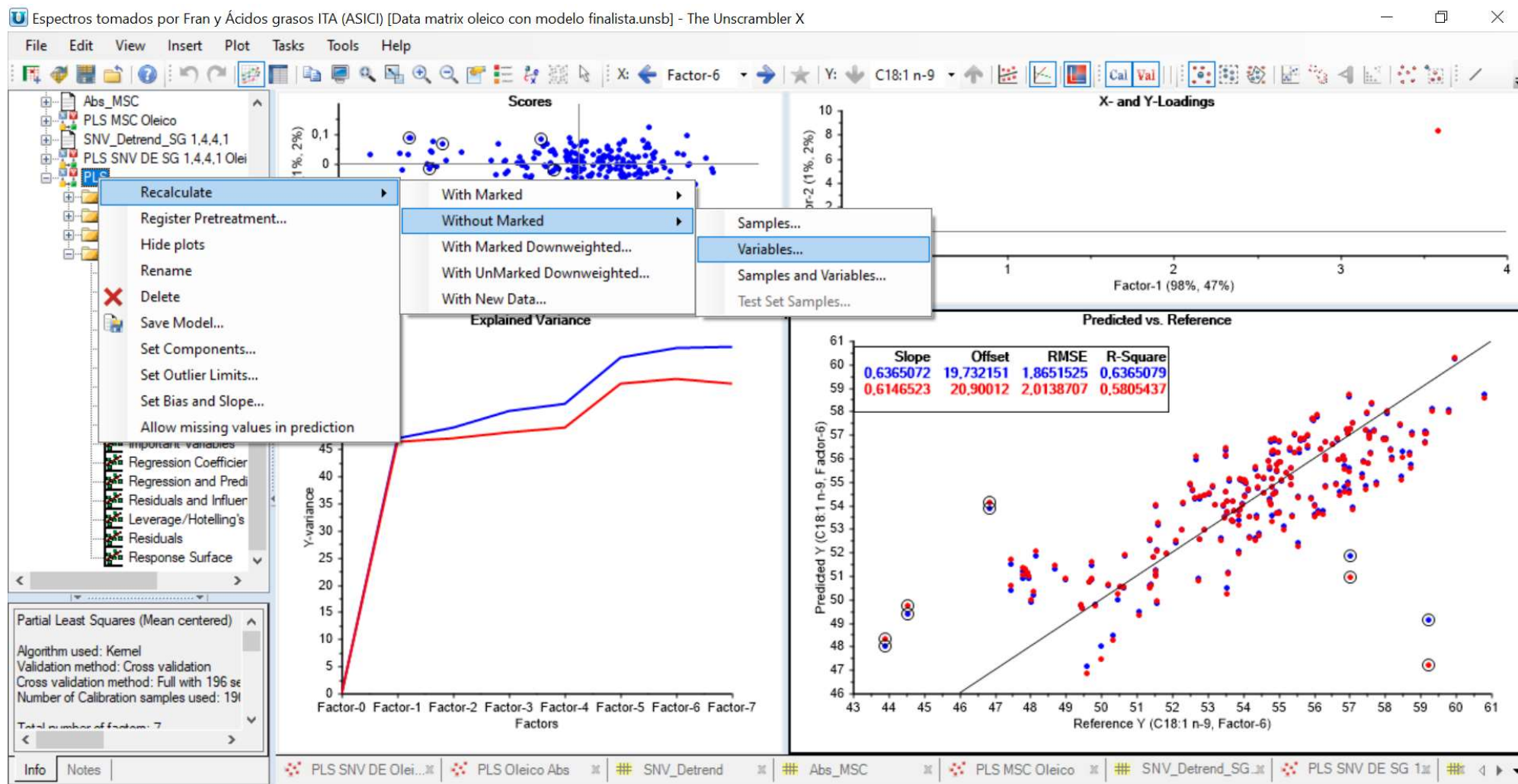
- Outliers. Recalculate the model without outliers

Marc one by one

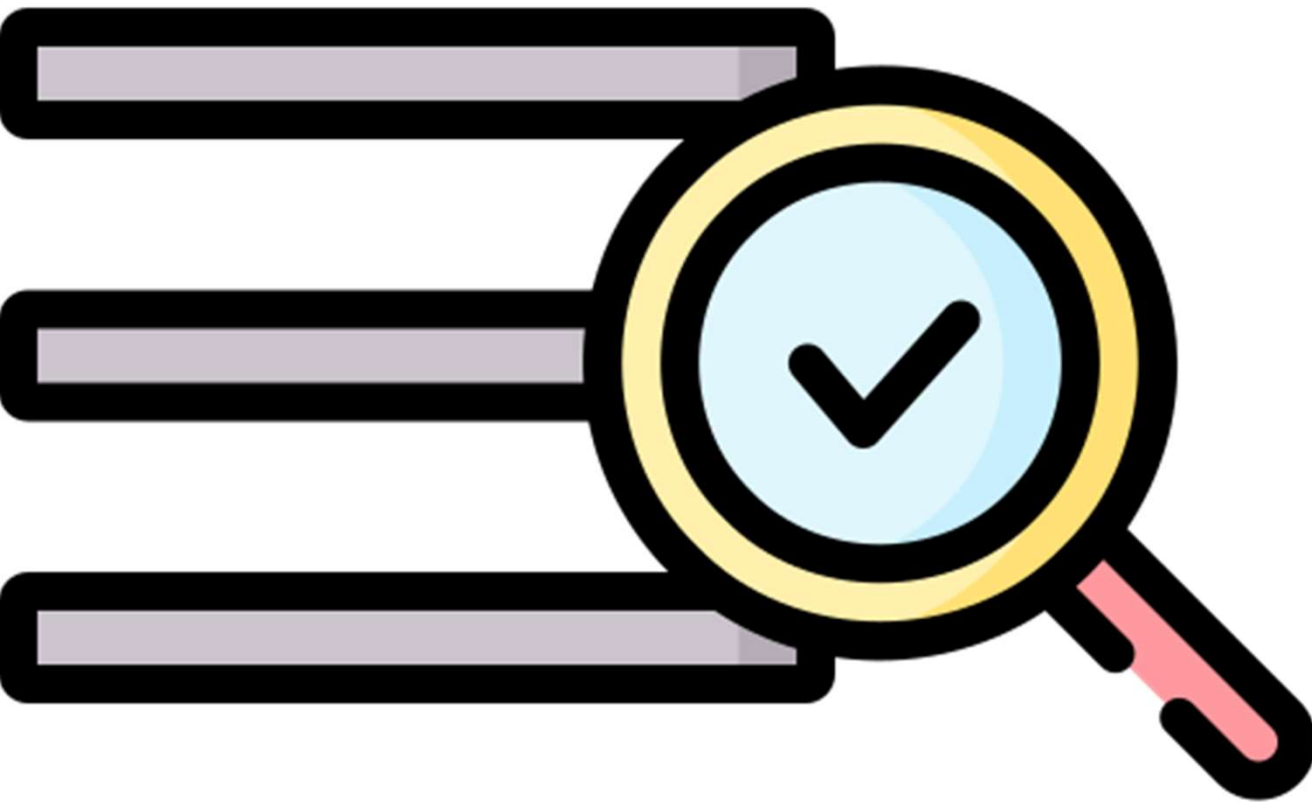




- Outliers. Recalculate the model without outliers







## Index

- What is NIRS technology?
- The Unscrambler
  - The construction of the matrix
  - Quantitative models
  - **Qualitative models**
  - Pre-treatments



## Qualitative models

- Data import
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)
- Development of classificatory models
- Predict



## Qualitative models

- **Data import**
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)
- Development of classificatory models
- Predict

- Data import

U TOTAL SPECTRA LONCHEADOS [Data matrix jamón (cualitativo).unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

U TOTAL SPECTRA LC

Matriz inicial  
Temperatura MAI  
Tiempo MAP Dur  
Matriz para trabajo

Matriz para tr		908.1	914.294	920.489	926.683	932.877	939.072	945.266	951.46	957.655	963.849	970.043	976.238	982.432	988.626	994.821	1001.015	1007.209
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Atmosf_Ctl_t	1	0,4679	0,4828	0,4984	0,4487	0,4073	0,3974	0,3961	0,3972	0,3983	0,4014	0,4050	0,4068	0,4074	0,4078	0,4076	0,4065	0,4
Atmosf_Ctl_t	2	0,4488	0,4636	0,4788	0,4281	0,3859	0,3759	0,3736	0,3734	0,3742	0,3773	0,3805	0,3820	0,3827	0,3831	0,3830	0,3819	0,3
Atmosf_Ctl_t	3	0,4585	0,4753	0,4917	0,4360	0,3903	0,3800	0,3784	0,3790	0,3815	0,3851	0,3886	0,3901	0,3910	0,3917	0,3911	0,3897	0,3
Atmosf_Ctl_t	4	0,4287	0,4451	0,4617	0,4074	0,3624	0,3515	0,3496	0,3496	0,3505	0,3536	0,3570	0,3583	0,3589	0,3593	0,3589	0,3573	0,3
Atmosf_Ctl_t	5	0,4347	0,4514	0,4687	0,4141	0,3689	0,3585	0,3570	0,3576	0,3590	0,3621	0,3656	0,3672	0,3678	0,3681	0,3672	0,3660	0,3
Atmosf_Ctl_t	6	0,4135	0,4311	0,4484	0,3933	0,3478	0,3363	0,3333	0,3331	0,3343	0,3374	0,3408	0,3424	0,3428	0,3427	0,3424	0,3410	0,3
Atmosf_Ctl_t	7	0,5057	0,5272	0,5477	0,4789	0,4236	0,4098	0,4067	0,4066	0,4075	0,4110	0,4146	0,4158	0,4163	0,4164	0,4160	0,4141	0,4
Atmosf_Ctl_t	8	0,4429	0,4606	0,4786	0,4220	0,3761	0,3653	0,3636	0,3638	0,3652	0,3689	0,3724	0,3741	0,3747	0,3749	0,3743	0,3727	0,3
Atmosf_Ctl_t	9	0,4748	0,4947	0,5146	0,4528	0,4029	0,3915	0,3897	0,3900	0,3917	0,3953	0,3990	0,4008	0,4015	0,4019	0,4018	0,4003	0,3
Atmosf_Ctl_t	10	0,4255	0,4450	0,4642	0,4037	0,3548	0,3424	0,3399	0,3398	0,3402	0,3430	0,3463	0,3479	0,3482	0,3485	0,3478	0,3466	0,3
Atmosf_Ctl_t	11	0,4913	0,5117	0,5318	0,4671	0,4154	0,4027	0,3999	0,3999	0,4011	0,4045	0,4081	0,4094	0,4098	0,4098	0,4093	0,4082	0,4
Atmosf_Ctl_t	12	0,5009	0,5211	0,5395	0,4803	0,4343	0,4243	0,4229	0,4238	0,4252	0,4288	0,4322	0,4340	0,4343	0,4348	0,4342	0,4331	0,4
Atmosf_Ctl_t	13	0,5305	0,5509	0,5703	0,5092	0,4613	0,4510	0,4499	0,4520	0,4546	0,4583	0,4623	0,4641	0,4645	0,4649	0,4646	0,4632	0,4
Atmosf_Ctl_t	14	0,5231	0,5445	0,5658	0,4998	0,4484	0,4362	0,4342	0,4345	0,4353	0,4389	0,4425	0,4441	0,4448	0,4452	0,4450	0,4433	0,4
Atmosf_Ctl_t	15	0,4609	0,4818	0,5015	0,4373	0,3873	0,3759	0,3732	0,3728	0,3731	0,3756	0,3795	0,3811	0,3816	0,3822	0,3819	0,3807	0,3
Atmosf_Ctl_t	16	0,4456	0,4659	0,4854	0,4242	0,3766	0,3651	0,3629	0,3628	0,3630	0,3654	0,3686	0,3699	0,3707	0,3711	0,3706	0,3693	0,3
Atmosf_Ctl_t	17	0,4652	0,4848	0,5032	0,4423	0,3957	0,3860	0,3847	0,3856	0,3874	0,3910	0,3947	0,3963	0,3972	0,3978	0,3979	0,3962	0,3
Atmosf_Ctl_t	18	0,4561	0,4772	0,4961	0,4345	0,3872	0,3785	0,3778	0,3795	0,3817	0,3856	0,3892	0,3911	0,3921	0,3926	0,3925	0,3911	0,3
Atmosf_Ctl_t	19	0,4433	0,4625	0,4802	0,4226	0,3791	0,3705	0,3696	0,3714	0,3736	0,3770	0,3806	0,3824	0,3834	0,3840	0,3839	0,3822	0,3
Atmosf_Ctl_t	20	0,4696	0,4887	0,5069	0,4480	0,4021	0,3925	0,3912	0,3917	0,3931	0,3960	0,3995	0,4010	0,4016	0,4017	0,4011	0,3997	0,3
Atmosf_Ctl_t	21	0,4560	0,4774	0,4963	0,4347	0,3871	0,3760	0,3738	0,3738	0,3747	0,3778	0,3813	0,3827	0,3834	0,3839	0,3834	0,3822	0,3
Atmosf_Ctl_t	22	0,4423	0,4644	0,4846	0,4226	0,3750	0,3635	0,3605	0,3603	0,3611	0,3633	0,3666	0,3680	0,3688	0,3690	0,3682	0,3674	0,3
Atmosf_Ctl_t	23	0,4436	0,4648	0,4839	0,4222	0,3750	0,3656	0,3645	0,3654	0,3668	0,3704	0,3738	0,3756	0,3764	0,3766	0,3765	0,3749	0,3
Atmosf_Ctl_t	24	0,4376	0,4577	0,4763	0,4174	0,3722	0,3622	0,3606	0,3608	0,3618	0,3647	0,3675	0,3688	0,3692	0,3694	0,3685	0,3667	0,3
Atmosf_Exp_	25	0,4721	0,4962	0,5189	0,4498	0,3979	0,3867	0,3854	0,3872	0,3897	0,3943	0,3985	0,4003	0,4007	0,4012	0,4002	0,3982	0,3
Atmosf_Exp_	26	0,4960	0,5181	0,5396	0,4727	0,4212	0,4114	0,4105	0,4122	0,4149	0,4189	0,4234	0,4257	0,4266	0,4271	0,4264	0,4249	0,4
Atmosf_Exp_	27	0,4735	0,4956	0,5166	0,4513	0,4018	0,3920	0,3911	0,3932	0,3958	0,4003	0,4046	0,4068	0,4077	0,4082	0,4077	0,4064	0,4
Atmosf_Exp_	28	0,4503	0,4723	0,4932	0,4300	0,3808	0,3703	0,3686	0,3691	0,3713	0,3755	0,3796	0,3812	0,3821	0,3824	0,3816	0,3798	0,3
Atmosf_Exp_	29	0,4629	0,4843	0,5056	0,4405	0,3912	0,3806	0,3793	0,3811	0,3836	0,3877	0,3919	0,3936	0,3945	0,3949	0,3939	0,3925	0,3
Atmosf_Exp_	30	0,4464	0,4678	0,4879	0,4243	0,3759	0,3663	0,3654	0,3676	0,3706	0,3751	0,3794	0,3811	0,3818	0,3824	0,3816	0,3798	0,3
Atmosf_Exp_	31	0,4563	0,4760	0,4945	0,4370	0,3928	0,3840	0,3833	0,3854	0,3882	0,3925	0,3964	0,3980	0,3987	0,3989	0,3985	0,3969	0,3
Atmosf_Exp_	32	0,4575	0,4779	0,4970	0,4378	0,3922	0,3839	0,3833	0,3854	0,3883	0,3924	0,3961	0,3980	0,3987	0,3989	0,3981	0,3965	0,3
Atmosf_Exp_	33	0,5693	0,5866	0,6028	0,5584	0,5236	0,5173	0,5173	0,5196	0,5218	0,5259	0,5288	0,5304	0,5315	0,5322	0,5324	0,5314	0,5
Atmosf_Exp_	34	0,4735	0,4958	0,5173	0,4520	0,4006	0,3880	0,3853	0,3850	0,3854	0,3883	0,3918	0,3930	0,3932	0,3929	0,3921	0,3908	0,3
Atmosf_Exp_	35	0,4621	0,4836	0,5033	0,4418	0,3936	0,3818	0,3795	0,3796	0,3808	0,3841	0,3874	0,3886	0,3887	0,3887	0,3876	0,3862	0,3
Atmosf_Exp_	36	0,4441	0,4641	0,4830	0,4233	0,3772	0,3672	0,3658	0,3670	0,3695	0,3734	0,3776	0,3791	0,3794	0,3795	0,3789	0,3771	0,3
Atmosf_Exp_	37	0,4629	0,4840	0,5035	0,4409	0,3923	0,3815	0,3801	0,3816	0,3837	0,3879	0,3916	0,3934	0,3944	0,3947	0,3941	0,3924	0,3
Atmosf_Exp_	38	0,4726	0,4940	0,5150	0,4472	0,3946	0,3836	0,3823	0,3842	0,3875	0,3920	0,3966	0,3987	0,3995	0,4000	0,3991	0,3974	0,3

Name : Matriz para trabajo  
Size : 133X125  
Created : 30/10/2024 12

Info Notes

Matriz inicial Matriz para trabajo

Ready Value: 0,4679017 133X125



## Qualitative models

- Data import
- **Insert row/column (to incorporate new samples or physical-chemical variables to predict)**
- **Insert Category variable**
- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)
- Development of classificatory models
- Predict



- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable

The following steps illustrate how to insert a category variable into the data matrix:

- Insert Menu:** In the 'TOTAL SPECTRA LONCHADOS' window, the 'Insert' menu is open, and 'Category Variable...' is selected under the 'Insert' option.
- Category Variable Dialog:** The 'Category Variable' dialog box is shown. The 'Category Name' is 'Time'. The 'Method' is 'Specify the level manually'. The 'Category Name' field is 'T5'. The 'Add' button is highlighted. The 'OK' button is also highlighted.
- Data Matrix:** The resulting data matrix is shown. A new column labeled 'Time' has been added to the matrix, with values 1, 2, 3, and 4 for the first four rows.



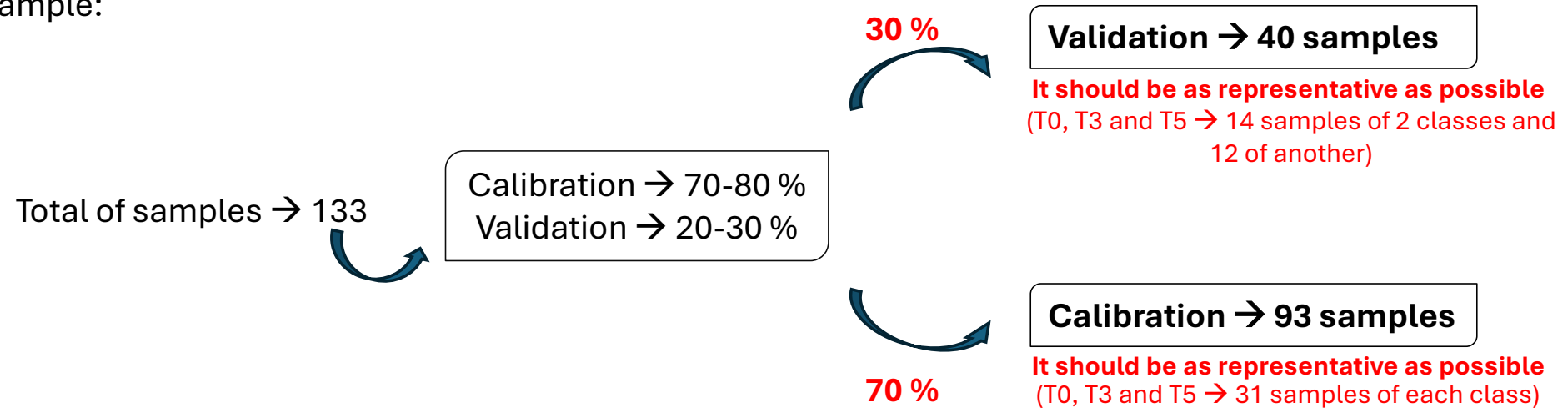


## Qualitative models

- Data import
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- **Define range for the purpose of grouping rows and columns**
- **Sets (Calibration and Validation)**
- Development of classificatory models
- Predict

- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)

Example:



Validation		
T0	T3	T5
14	14	12



- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)

The screenshot shows the 'TOTAL SPECTRA LONCHEADOS' data matrix in The Unscrambler X. The 'Matriz para trabajar' is selected, and the 'Range' menu is open, highlighting 'Define Range...'. A blue arrow points from this menu item to the 'Define Range' dialog box on the right.

The 'Define Range' dialog box has the following fields:

- Row ranges:** All spectra (2-126), All spectra acotad (17-126)
- Column ranges:** (Empty)
- Rowset:** 9-10 (highlighted with a red circle and a yellow starburst)
- Columnset:** (Empty)
- Keep out:** Rows (Empty), Columns (Empty)
- Special intervals:** (Collapsed)
- Buttons:** Update, Create (highlighted with a red circle), OK (highlighted with a red circle), Cancel

The data matrix 'Matriz para trabajar' is displayed below the dialog box. It has 13 columns: Time, 908.1, 914.294, 920.489, 926.683, 932.877, 939.072, 945.266, 951.46. The rows are grouped by 'Time' (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13). The 'Time' column is highlighted in orange, and the '908.1' column is highlighted in blue. The 'Create' button in the dialog box is highlighted with a red circle and a yellow starburst.

Matriz para trabajar	Time	908.1	914.294	920.489	926.683	932.877	939.072	945.266	951.46
Atmosf_Ctl_t0_1_0902	1	0,4679	0,4828	0,4984	0,4487	0,4073	0,3974	0,3961	0,
Atmosf_Ctl_t0_2_0902	2	0,4488	0,4636	0,4788	0,4281	0,3859	0,3759	0,3736	0,
Atmosf_Ctl_t0_3_0902	3	0,4585	0,4753	0,4917	0,4360	0,3903	0,3800	0,3784	0,
Atmosf_Ctl_t0_4_0902	4	0,4287	0,4451	0,4617	0,4074	0,3624	0,3515	0,3496	0,
Atmosf_Ctl_t0_5_0902	5	0,4347	0,4514	0,4687	0,4141	0,3689	0,3585	0,3570	0,
Atmosf_Ctl_t0_6_0902	6	0,4135	0,4311	0,4484	0,3933	0,3478	0,3363	0,3333	0,
Atmosf_Ctl_t0_7_0902	7	0,5057	0,5272	0,5477	0,4789	0,4236	0,4098	0,4067	0,
Atmosf_Ctl_t0_8_0902	8	0,4429	0,4606	0,4786	0,4220	0,3761	0,3653	0,3636	0,
Atmosf_Ctl_t0_9_0902	9	0,4748	0,4947	0,5146	0,4528	0,4029	0,3915	0,3897	0,
Atmosf_Ctl_t0_10_090	10	0,4255	0,4450	0,4642	0,4037	0,3548	0,3424	0,3399	0,
Atmosf_Ctl_t0_11_090	11	0,4913	0,5117	0,5318	0,4671	0,4154	0,4027	0,3999	0,
Atmosf_Ctl_t0_12_090	12	0,5009	0,5211	0,5395	0,4803	0,4343	0,4243	0,4229	0,
Atmosf_Ctl_t0_13_090	13	0,5305	0,5509	0,5703	0,5092	0,4613	0,4510	0,4499	0,





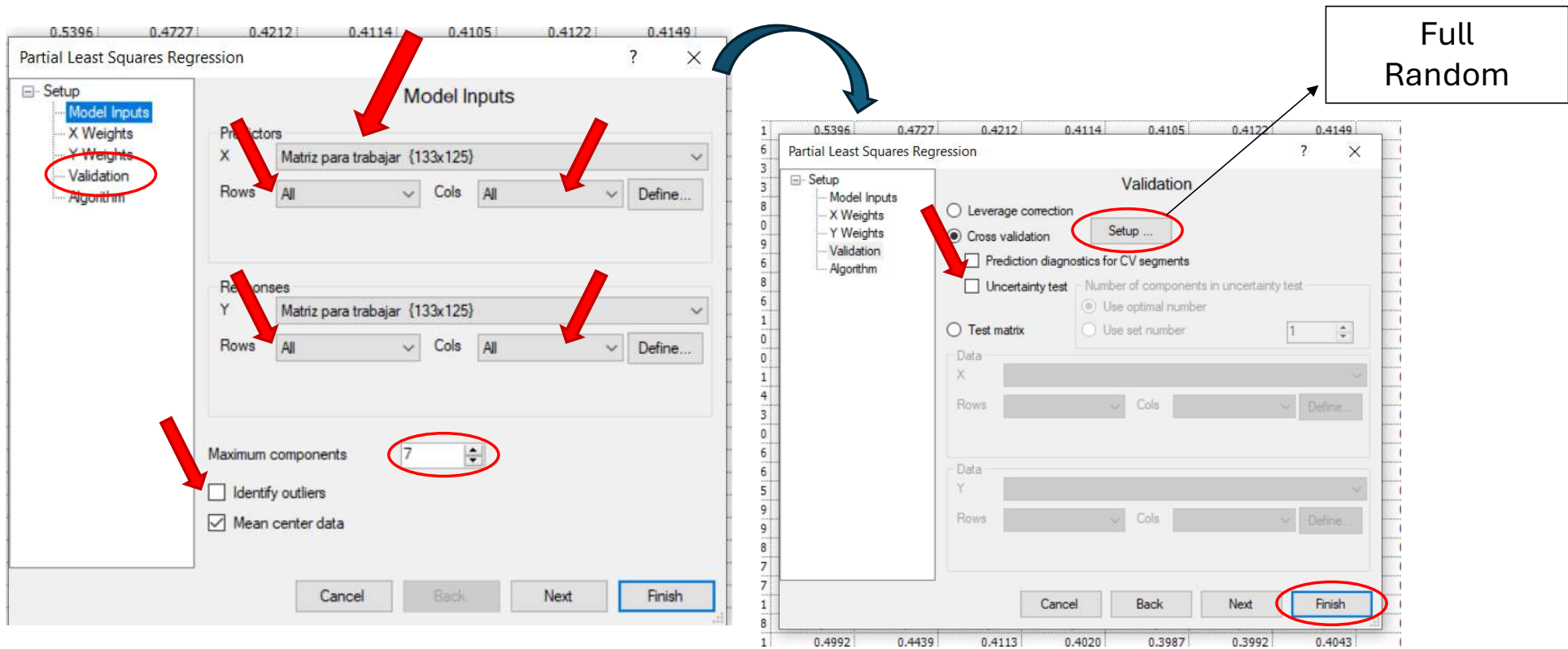
## Qualitative models

- Data import
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)
- **Development of classificatory models**
- Predict

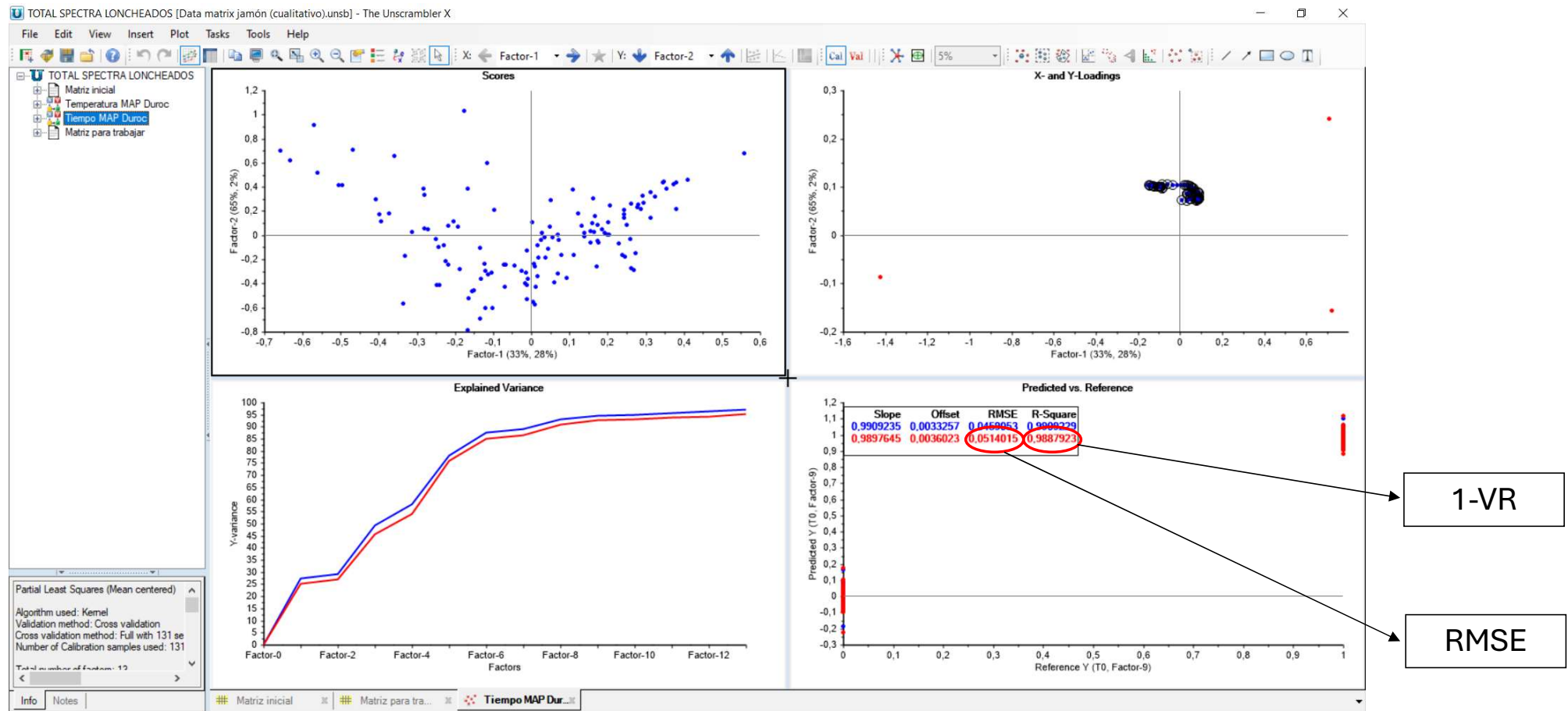
- Development of classificatory models

The screenshot displays the Unscrambler X software interface. The 'Analyze' menu is open, showing various statistical methods, with 'Partial Least Squares Regression...' selected. The PLS Regression dialog box is open, showing the 'Setup' tab. The 'Model Inputs' section has 'Matriz para trabajar {133x125}' selected for both X and Y matrices. The 'Maximum components' is set to 7, and 'Mean center data' is checked. Red arrows point to the matrix selection, the maximum components field, and the 'Mean center data' checkbox.

- Development of classificatory models

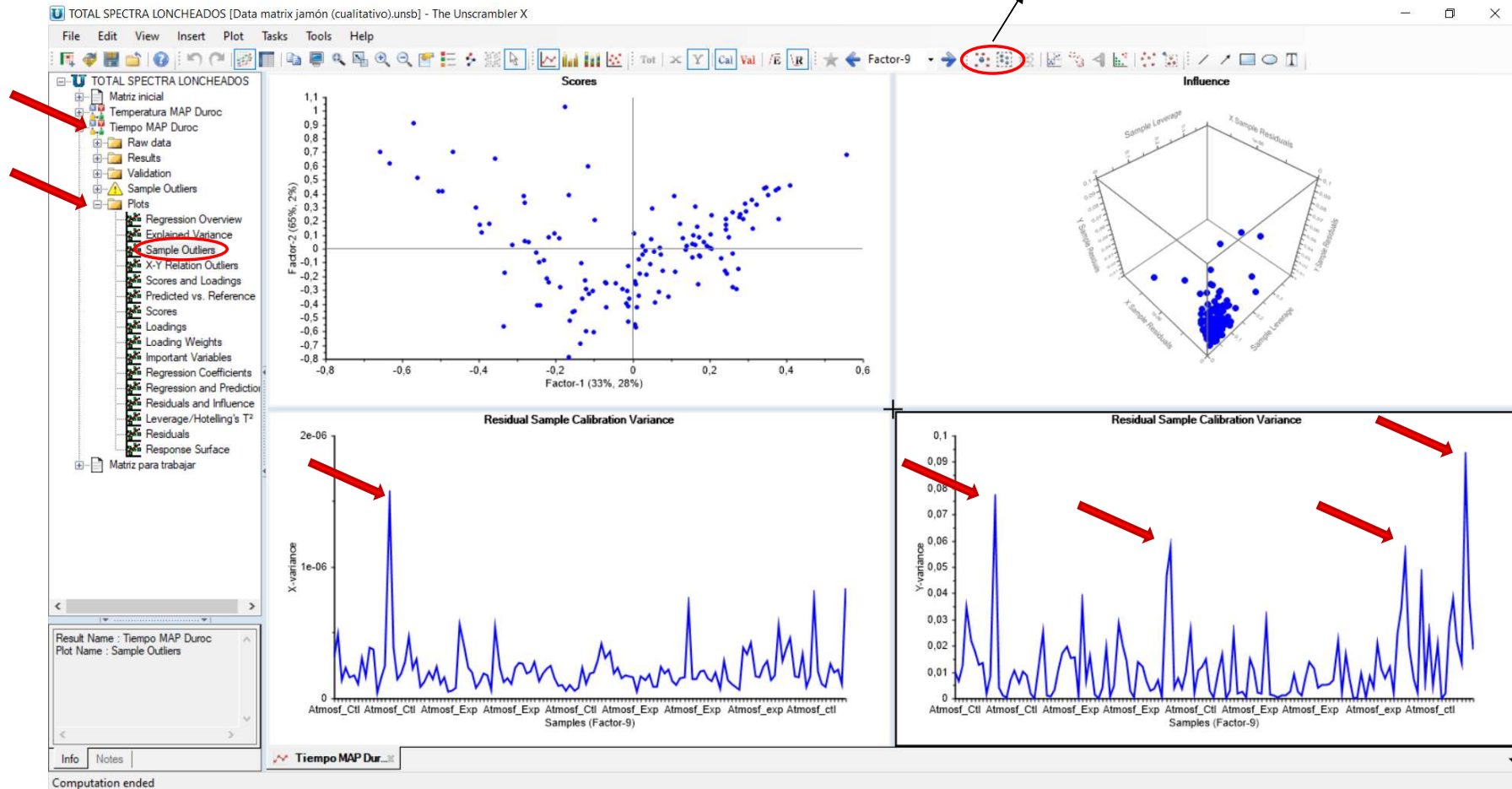


- Development of classificatory models



- Development of classificatory models

Click to select the outliers we have identified







## Qualitative models

- Data import
- Insert row/column (to incorporate new samples or physical-chemical variables to predict)
- Insert Category variable
- Define range for the purpose of grouping rows and columns
- Sets (Calibration and Validation)
- Development of classificatory models
- **Predict**

- Predict

## STATISTICS

### **SENSITIVITY (SE)**

Which denotes the percentage of samples belonging to an established class that the studied model has recognized as belonging to that class

$$SE (\%) = \frac{TP}{TP + FN}$$

TP = True positive  
FN = False negative

### **SPECIFICITY (SP)**

Which refers to the percentage of samples that do not belong to the target class and that the model has correctly rejected

$$SP (\%) = \frac{TN}{TN + FP}$$

TN = True negative  
FP = False positive

- Predict

## STATISTICS

### ACCURACY


Defined as the percentage of samples correctly classified during external validation

$$Accuracy (\%) = \frac{TN + TP}{TN + TP + FN + FP}$$

TP = True positive  
TN = True negative  
FP = False positive  
FN = False negative

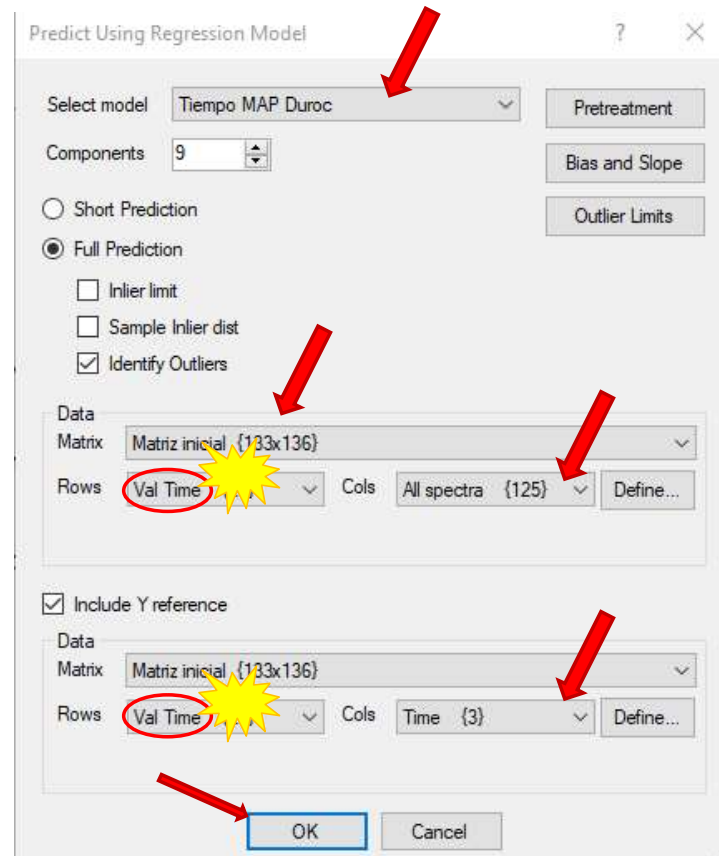
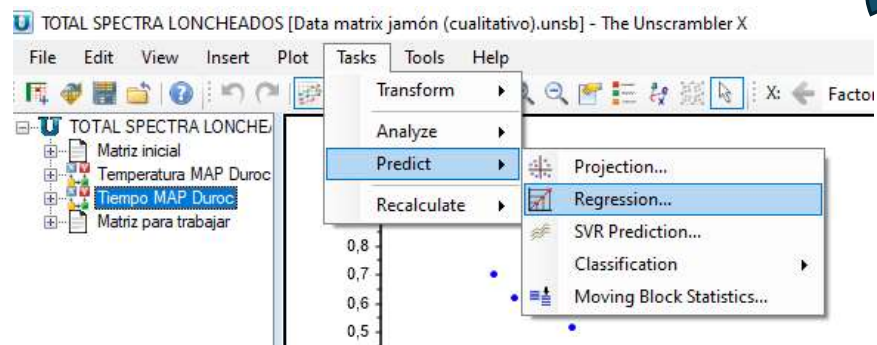
### MATTHEWS CORRELATION COEFFICIENT (MCC)

Which is suitable for assessing the performance of the model when the number of samples in the various classes is not the same

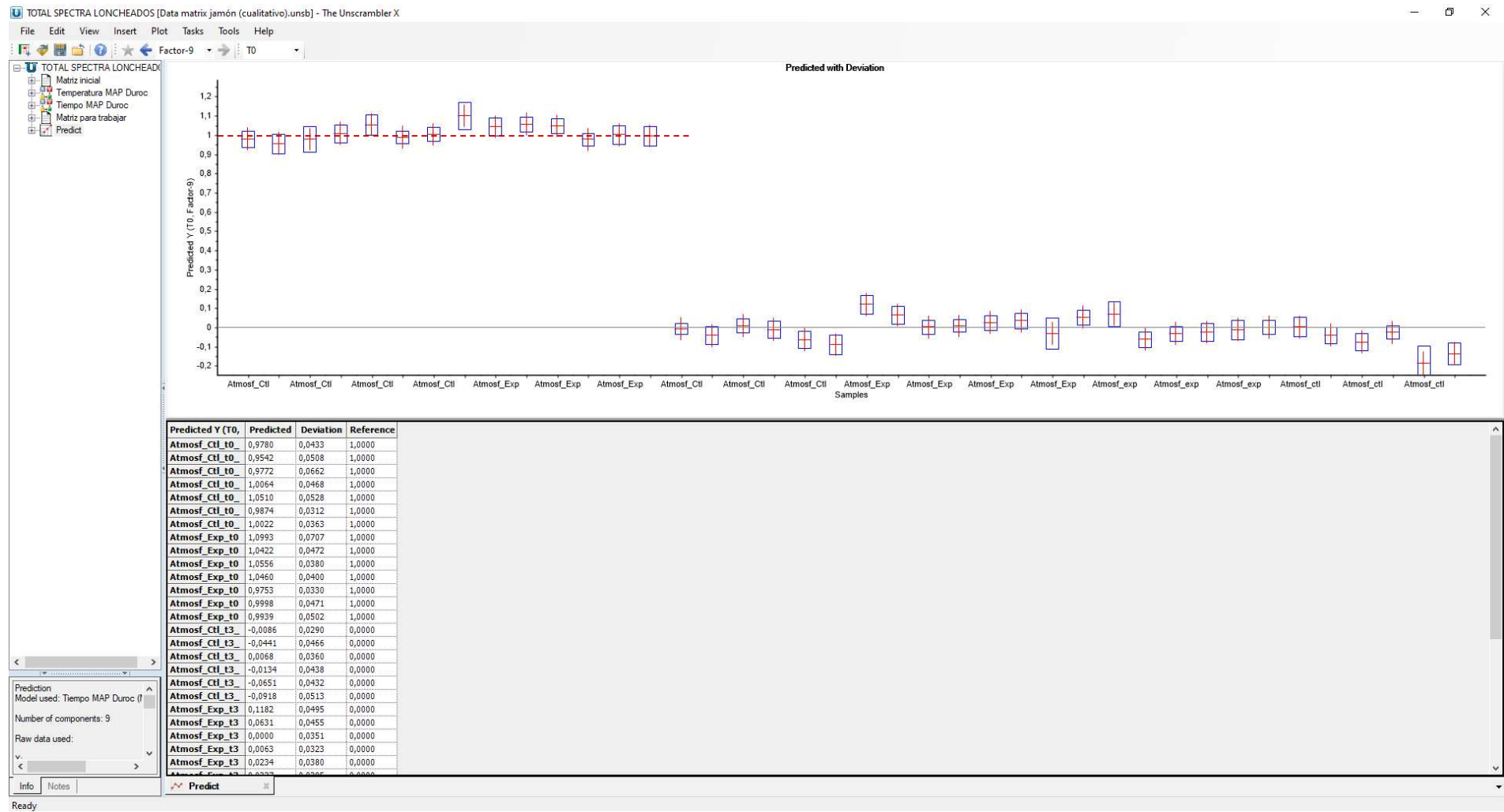

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FN)(TN + TP)}}$$

Value of 1 indicates a perfect prediction; -1 represents a total disagreement between the prediction and the actual values; and 0 means “no better than a random prediction”

- Predict

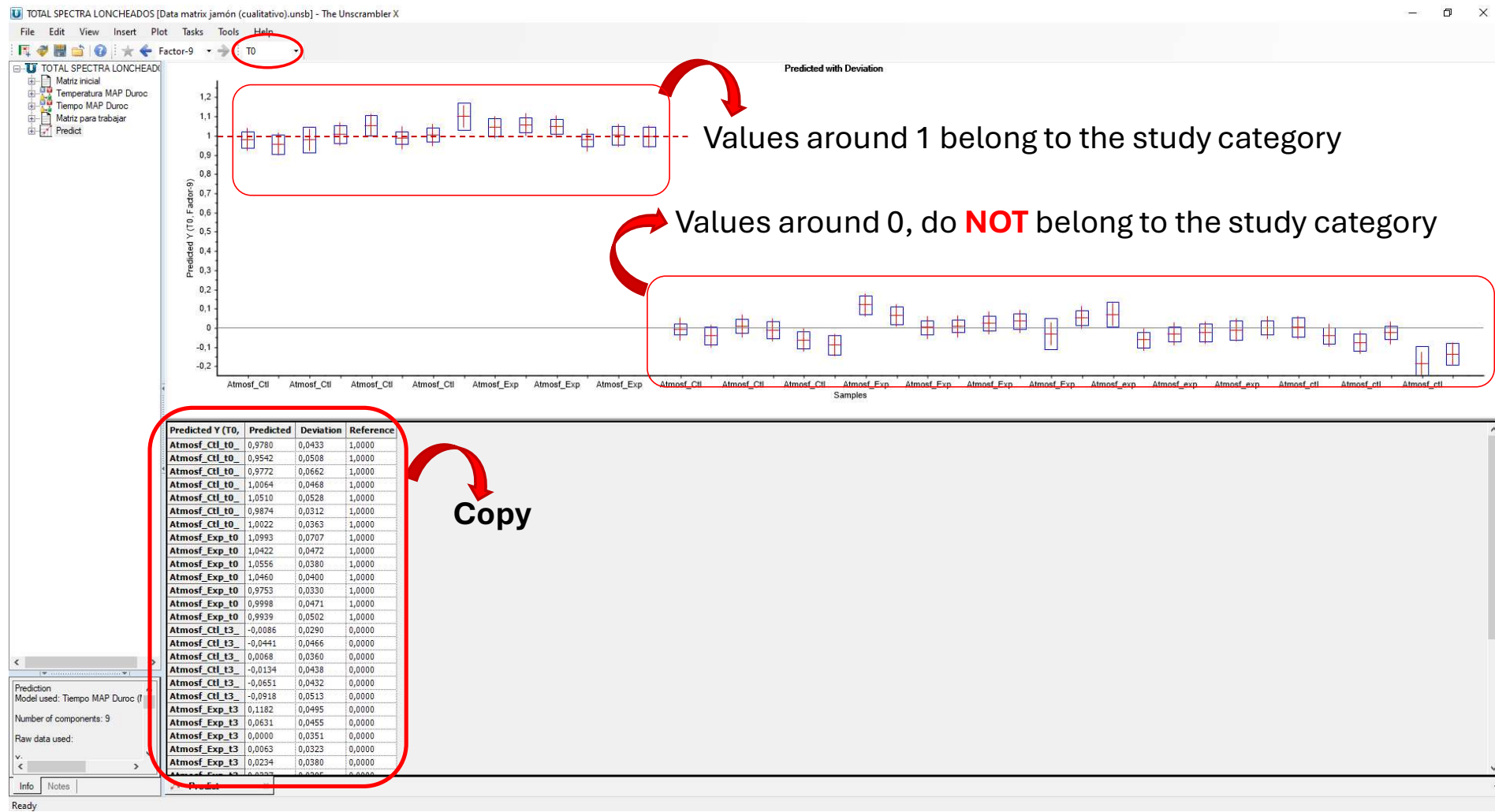


- Predict



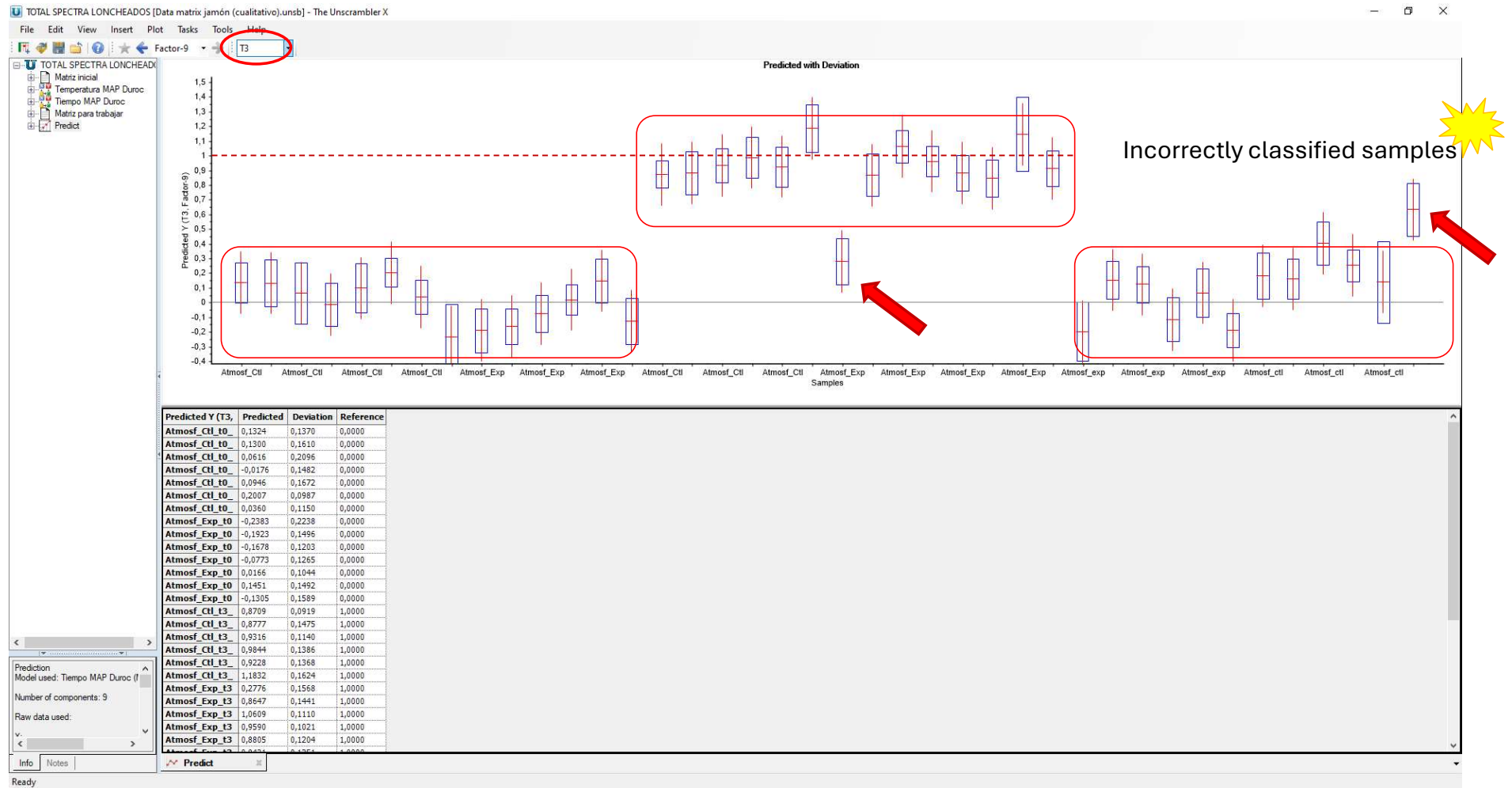


- Predict



- Predict

Another example:

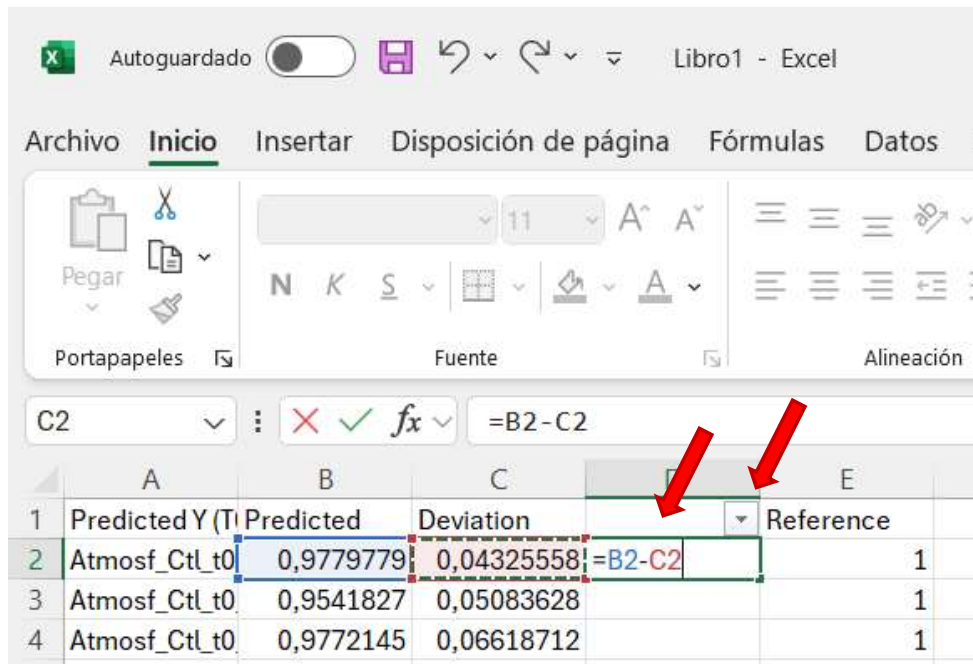


- Predict

Predicted value - deviation  $\geq 0.5 \rightarrow$  **SENSITIVITY**.

Predicted value + deviation  $< 0.5 \rightarrow$  **SPECIFICITY**.

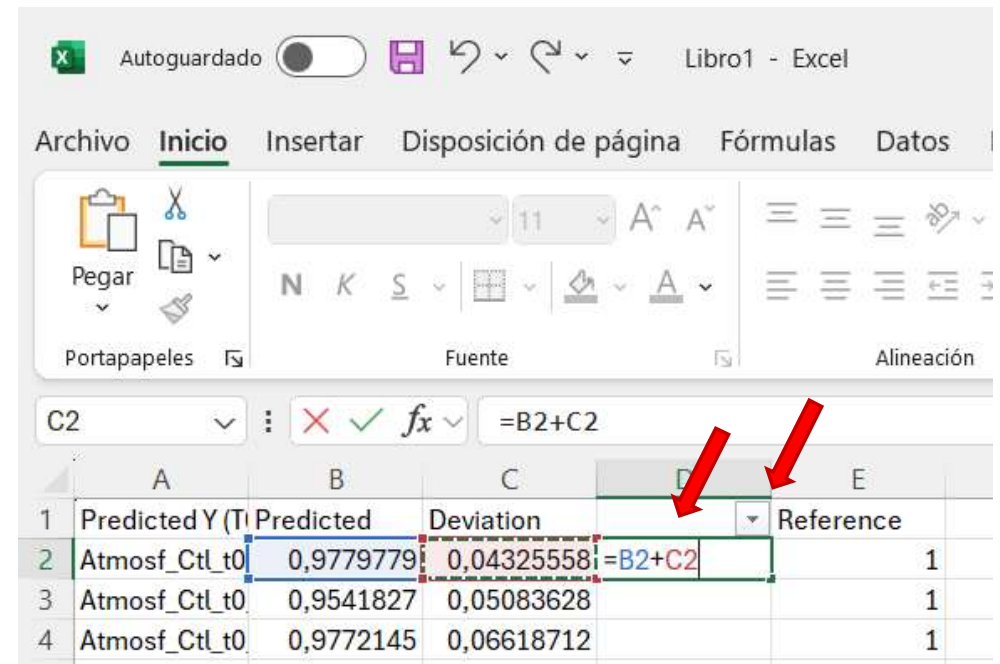
**We calculate the most unfavorable case**



This screenshot shows an Excel spreadsheet with the formula bar set to `=B2-C2`. The spreadsheet has columns A (Predicted Y (T)), B (Predicted), C (Deviation), and E (Reference). Row 2 is highlighted, showing a predicted value of 0.9779779 and a deviation of 0.04325558. Red arrows point to the deviation cell (C2) and the formula bar.

	A	B	C	E
1	Predicted Y (T)	Predicted	Deviation	Reference
2	Atmosf_Ctl_t0	0,9779779	0,04325558	1
3	Atmosf_Ctl_t0	0,9541827	0,05083628	1
4	Atmosf_Ctl_t0	0,9772145	0,06618712	1

**SENSITIVITY**



This screenshot shows an Excel spreadsheet with the formula bar set to `=B2+C2`. The spreadsheet has columns A (Predicted Y (T)), B (Predicted), C (Deviation), and E (Reference). Row 2 is highlighted, showing a predicted value of 0.9779779 and a deviation of 0.04325558. Red arrows point to the deviation cell (C2) and the formula bar.

	A	B	C	E
1	Predicted Y (T)	Predicted	Deviation	Reference
2	Atmosf_Ctl_t0	0,9779779	0,04325558	1
3	Atmosf_Ctl_t0	0,9541827	0,05083628	1
4	Atmosf_Ctl_t0	0,9772145	0,06618712	1

**SPECIFICITY**

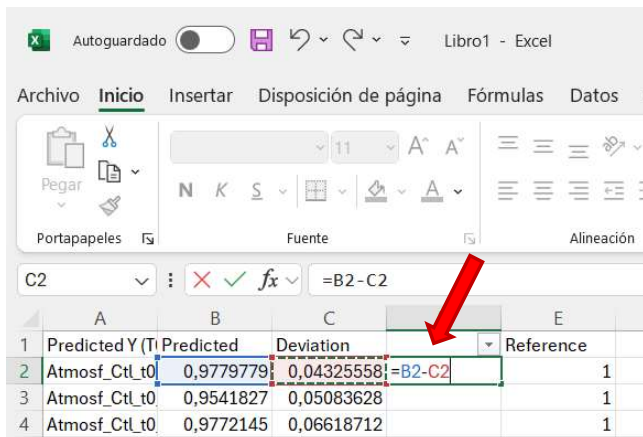
- Predict

Predicted value - deviation  $\geq 0.5 \rightarrow$  **SENSITIVITY**.

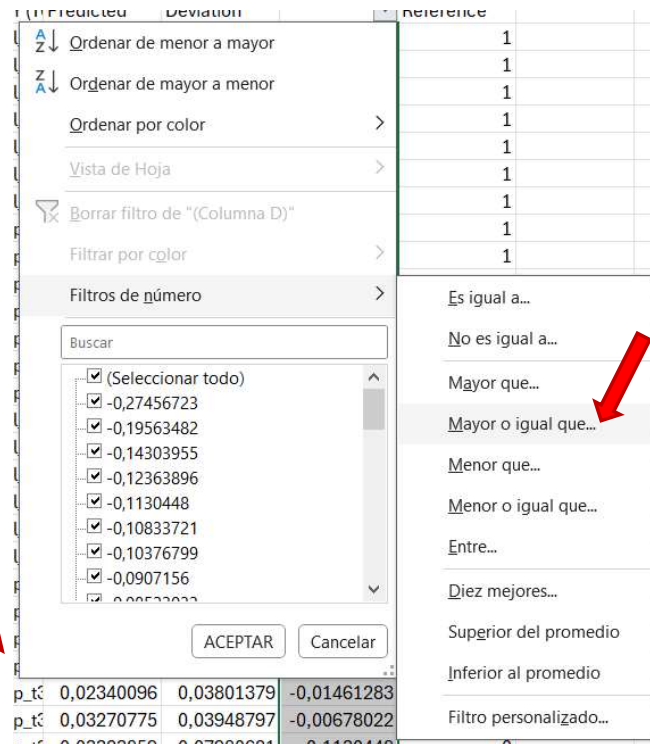
Predicted value + deviation  $< 0.5 \rightarrow$  **SPECIFICITY**.

**We calculate the most unfavorable case**

## SENSITIVITY



	A	B	C	E
1	Predicted Y (T)	Predicted	Deviation	Reference
2	Atmosf_Ctl_t0	0,9779779	0,04325558	1
3	Atmosf_Ctl_t0	0,9541827	0,05083628	1
4	Atmosf_Ctl_t0	0,9772145	0,06618712	1



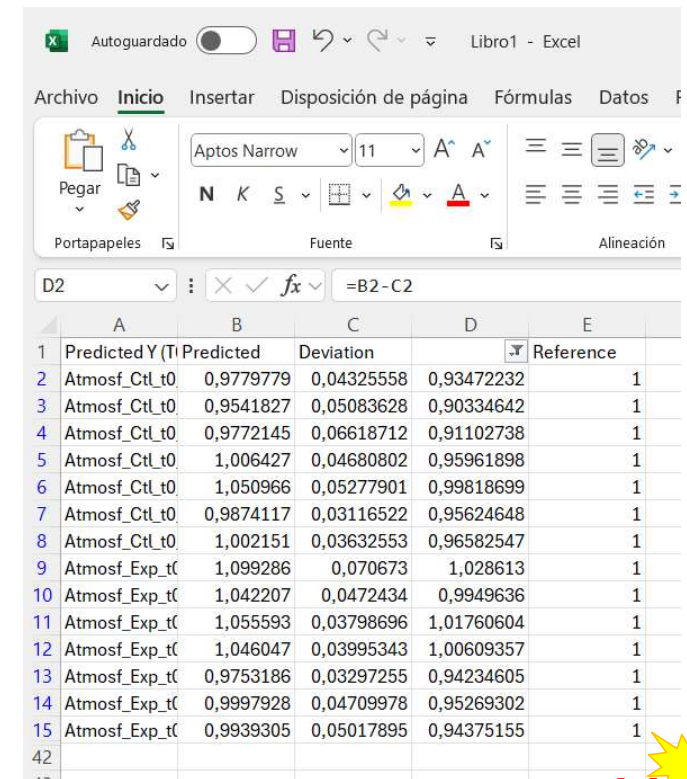
Ordenar de menor a mayor  
Ordenar de mayor a menor  
Ordenar por color  
Vista de Hoja  
Borrar filtro de "(Columna D)"  
Filtrar por color  
Filtros de número

Buscar

- ☒ (Seleccionar todo)
- ☒ -0,27456723
- ☒ -0,19563482
- ☒ -0,14303955
- ☒ -0,12363896
- ☒ -0,1130448
- ☒ -0,10833721
- ☒ -0,10376799
- ☒ -0,0907156
- ☒ 0,04325558

ACEPTAR Cancelar

Es igual a...  
No es igual a...  
Mayor que...  
**Mayor o igual que...**  
Menor que...  
Menor o igual que...  
Entre...  
Diez mejores...  
Superior del promedio  
Inferior al promedio  
Filtro personalizado...



	A	B	C	D	E
1	Predicted Y (T)	Predicted	Deviation	Reference	
2	Atmosf_Ctl_t0	0,9779779	0,04325558	0,93472232	1
3	Atmosf_Ctl_t0	0,9541827	0,05083628	0,90334642	1
4	Atmosf_Ctl_t0	0,9772145	0,06618712	0,91102738	1
5	Atmosf_Ctl_t0	1,006427	0,04680802	0,95961898	1
6	Atmosf_Ctl_t0	1,050966	0,05277901	0,99818699	1
7	Atmosf_Ctl_t0	0,9874117	0,03116522	0,95624648	1
8	Atmosf_Ctl_t0	1,002151	0,03632553	0,96582547	1
9	Atmosf_Exp_t0	1,099286	0,070673	1,028613	1
10	Atmosf_Exp_t0	1,042207	0,0472434	0,9949636	1
11	Atmosf_Exp_t0	1,055593	0,03798696	1,01760604	1
12	Atmosf_Exp_t0	1,046047	0,03995343	1,00609357	1
13	Atmosf_Exp_t0	0,9753186	0,03297255	0,94234605	1
14	Atmosf_Exp_t0	0,9997928	0,04709978	0,95269302	1
15	Atmosf_Exp_t0	0,9939305	0,05017895	0,94375155	1

14

True positive

- Predict

Validation		
T0	T3	T5
14	14	12

SF

14  
True positive

$$SE (\%) = \frac{TP}{TP + FN}$$

TP = True positive  
FN = False negative

FN = Total class samples - TP

$$SE (\%) = \frac{14}{14 + 0}$$

100%

TP = 14  
FN = 0



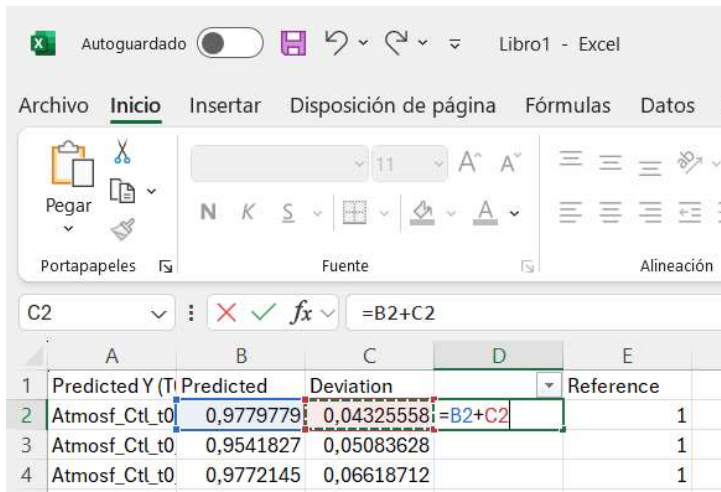
- Predict

Predicted value - deviation  $\geq 0.5 \rightarrow$  **SENSITIVITY.**

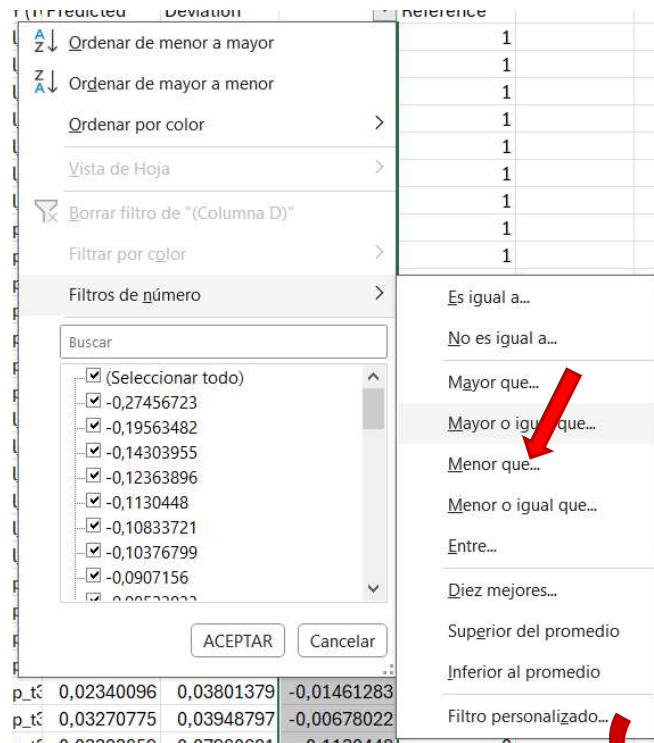
Predicted value + deviation  $< 0.5 \rightarrow$  **SPECIFICITY.**

**We calculate the most unfavorable case**

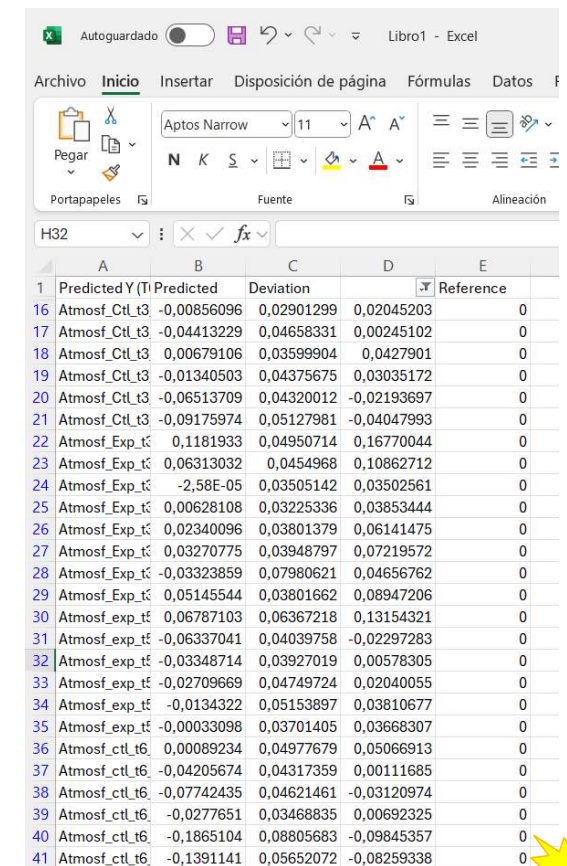
## SPECIFICITY



	A	B	C	D	E
1	Predicted Y (Ti Predicted	Deviation		Reference	
2	Atmosf_Ctl_t0	0,9779779	0,04325558	=B2+C2	1
3	Atmosf_Ctl_t0	0,9541827	0,05083628		1
4	Atmosf_Ctl_t0	0,9772145	0,06618712		1



Predicted	Deviation	Reference
1		1
1		1
1		1
1		1
1		1
1		1
1		1
1		1



	A	B	C	D	E
1	Predicted Y (Ti Predicted	Deviation		Reference	
16	Atmosf_Ctl_t3	-0,00856096	0,02901299	0,02045203	0
17	Atmosf_Ctl_t3	-0,04413229	0,04658331	0,00245102	0
18	Atmosf_Ctl_t3	0,00679106	0,03599904	0,0427901	0
19	Atmosf_Ctl_t3	-0,01340503	0,04375675	0,03035172	0
20	Atmosf_Ctl_t3	-0,06513709	0,04320012	-0,02193697	0
21	Atmosf_Ctl_t3	-0,09175974	0,05127981	-0,04047993	0
22	Atmosf_Exp_t3	0,1181933	0,04950714	0,16770044	0
23	Atmosf_Exp_t3	0,06313032	0,0454968	0,10862712	0
24	Atmosf_Exp_t3	-2,58E-05	0,03505142	0,03502561	0
25	Atmosf_Exp_t3	0,00628108	0,03225336	0,03853444	0
26	Atmosf_Exp_t3	0,02340096	0,03801379	0,06141475	0
27	Atmosf_Exp_t3	0,03270775	0,03948797	0,07219572	0
28	Atmosf_Exp_t3	-0,03323859	0,07980621	0,04656762	0
29	Atmosf_Exp_t3	0,05145544	0,03801662	0,08947206	0
30	Atmosf_Exp_t3	0,06787103	0,06367218	0,13154321	0
31	Atmosf_Exp_t3	-0,06337041	0,04039758	-0,02297283	0
32	Atmosf_Exp_t3	-0,03348714	0,03927019	0,00578305	0
33	Atmosf_Exp_t3	-0,02709669	0,04749724	0,02040055	0
34	Atmosf_Exp_t3	-0,0134322	0,05153897	0,03810677	0
35	Atmosf_Exp_t3	-0,00033098	0,03701405	0,03668307	0
36	Atmosf_Ctl_t6	0,00089234	0,04977679	0,05066913	0
37	Atmosf_Ctl_t6	-0,04205674	0,04317359	0,00111685	0
38	Atmosf_Ctl_t6	-0,07742435	0,04621461	-0,03120974	0
39	Atmosf_Ctl_t6	-0,0277651	0,03468835	0,00692325	0
40	Atmosf_Ctl_t6	-0,1865104	0,08805683	-0,09845357	0
41	Atmosf_Ctl_t6	-0,1391141	0,05652072	-0,08259338	0

26

True negative

- Predict

	A	B	C	D	E
1	Predicted Y (T)	Predicted	Deviation		Reference
16	Atmosf_Ctl_t3	-0,00856096	0,02901299	0,02045203	0
17	Atmosf_Ctl_t3	-0,04413229	0,04658331	0,00245102	0
18	Atmosf_Ctl_t3	0,00679106	0,03599904	0,0427901	0
19	Atmosf_Ctl_t3	-0,01340503	0,04375675	0,03035172	0
20	Atmosf_Ctl_t3	-0,06513709	0,04320012	-0,02193697	0
21	Atmosf_Ctl_t3	-0,09175974	0,05127981	-0,04047993	0
22	Atmosf_Exp_t3	0,1181933	0,04950714	0,16770044	0
23	Atmosf_Exp_t3	0,06313032	0,0454968	0,10862712	0
24	Atmosf_Exp_t3	-2,58E-05	0,03505142	0,03502561	0
25	Atmosf_Exp_t3	0,00628108	0,03225336	0,03853444	0
26	Atmosf_Exp_t3	0,02340096	0,03801379	0,06141475	0
27	Atmosf_Exp_t3	0,03270775	0,03948797	0,07219572	0
28	Atmosf_Exp_t3	-0,03323859	0,07980621	0,04656762	0
29	Atmosf_Exp_t3	0,05145544	0,03801662	0,08947206	0
30	Atmosf_exp_t6	0,06787103	0,06367218	0,13154321	0
31	Atmosf_exp_t6	-0,06337041	0,04039758	-0,02297283	0
32	Atmosf_exp_t6	-0,03348714	0,03927019	0,00578305	0
33	Atmosf_exp_t6	-0,02709669	0,04749724	0,02040055	0
34	Atmosf_exp_t6	-0,0134322	0,05153897	0,03810677	0
35	Atmosf_exp_t6	-0,00033098	0,03701405	0,03668307	0
36	Atmosf_ctl_t6	0,00089234	0,04977679	0,05066913	0
37	Atmosf_ctl_t6	-0,04205674	0,04317359	0,00111685	0
38	Atmosf_ctl_t6	-0,07742435	0,04621461	-0,03120974	0
39	Atmosf_ctl_t6	-0,0277651	0,03468835	0,00692325	0
40	Atmosf_ctl_t6	-0,1865104	0,08805683	-0,09845357	0
41	Atmosf_ctl_t6	-0,1391141	0,05652072	-0,08259338	0

**26**  
True negative

Validation		
T0	T3	T5
14	14	12

**26**

$$SP (\%) = \frac{TN}{TN + FP}$$

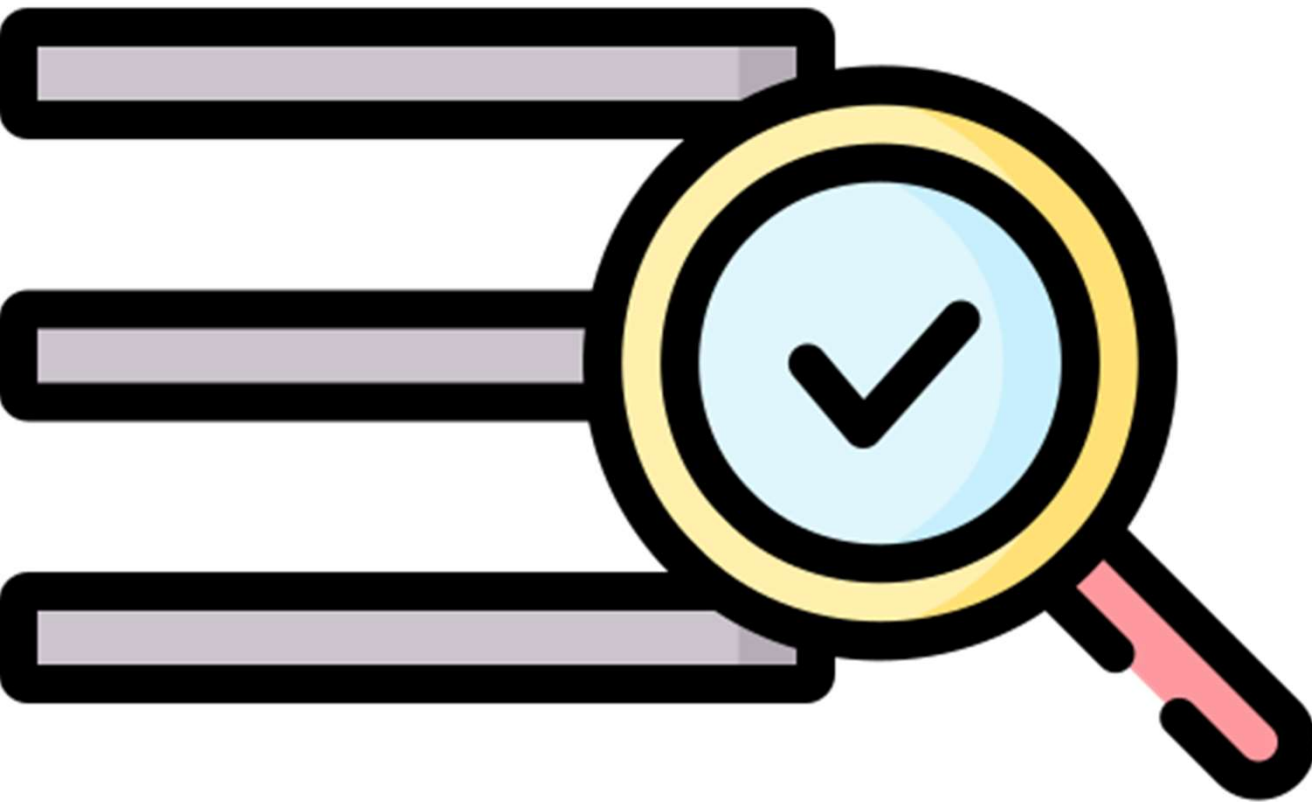
$$SP (\%) = \frac{26}{26 + 0}$$

**100%**

TN = True negative  
FP = False positive

TN = 26  
FP = 0

FP = Total of samples not belonging to the class - TN



## Index

- What is NIRS technology?
- The Unscrambler
  - The construction of the matrix
  - Quantitative models
  - Qualitative models
  - **Pre-treatments**

- Pre-treatments

Los **efectos no deseados presentes en las señales** obtenidas en espectroscopia se denominan comúnmente **ruido**.

Este ruido puede tener diferentes causas u orígenes.

- ☐ Puede haber ruido debido a los componentes de la instrumentación utilizada para el registro del espectro (ruido instrumental) (variaciones de temperatura, humedad u otras)
- ☐ o bien variaciones en la señal debidas a la propia naturaleza de la muestra



### **Pre-treatments**

Tratamientos espectrales para reducir el ruido

- Pre-treatments

- ☐ **Scattering effect** refers to the **dispersion of incident radiation** due to **physical phenomena** (texture, size and geometry of the particles that make up the sample). This causes unwanted variations in the spectral data.



#### Scatter correction treatments

- ☐ **Detrending (DT)**
- ☐ **Standard Normal Variate (SNV)**
- ☐ **Multiplicative Scatter Correction (MSC)**

- ☐ **Average**

Averaging several analytical signals of the same sample decreases noise and improves the signal-to-noise ratio.

- ☐ **Spectral smoothing** to **mathematically** reduce random noise and increase signal-to-noise ratio



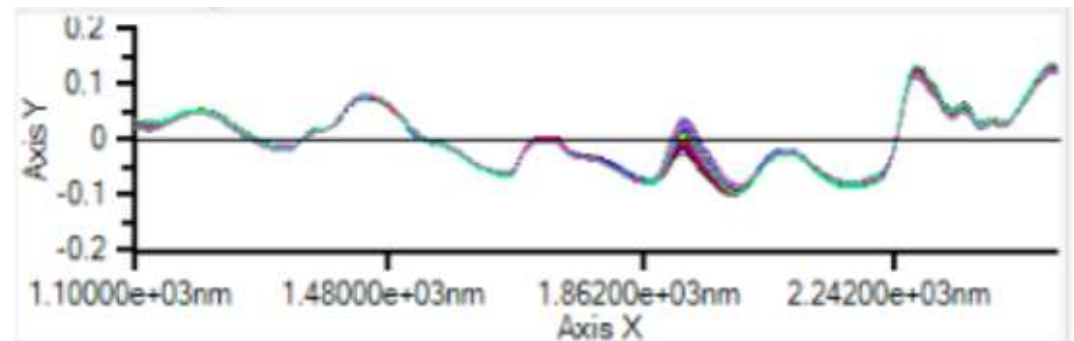
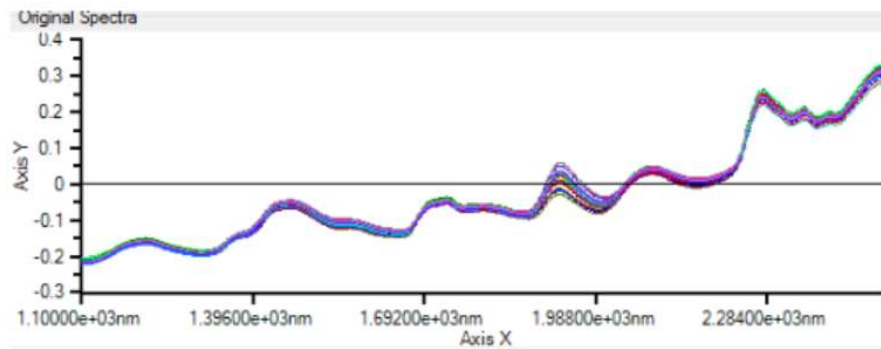
- ☐ **Derivatives**



- Pre-treatments

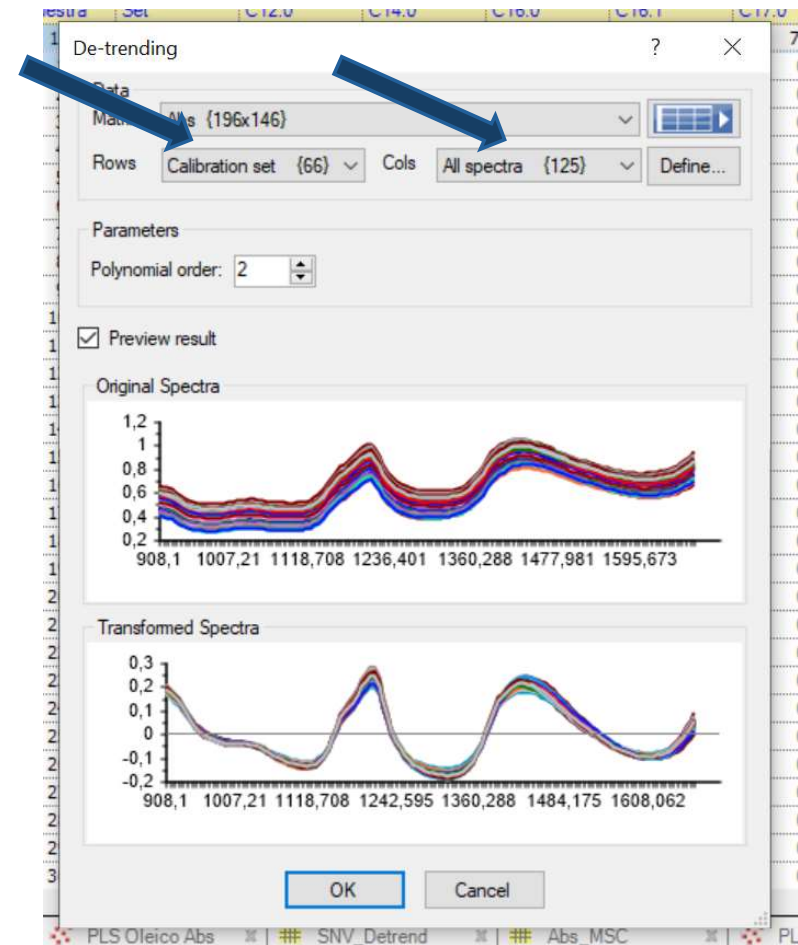
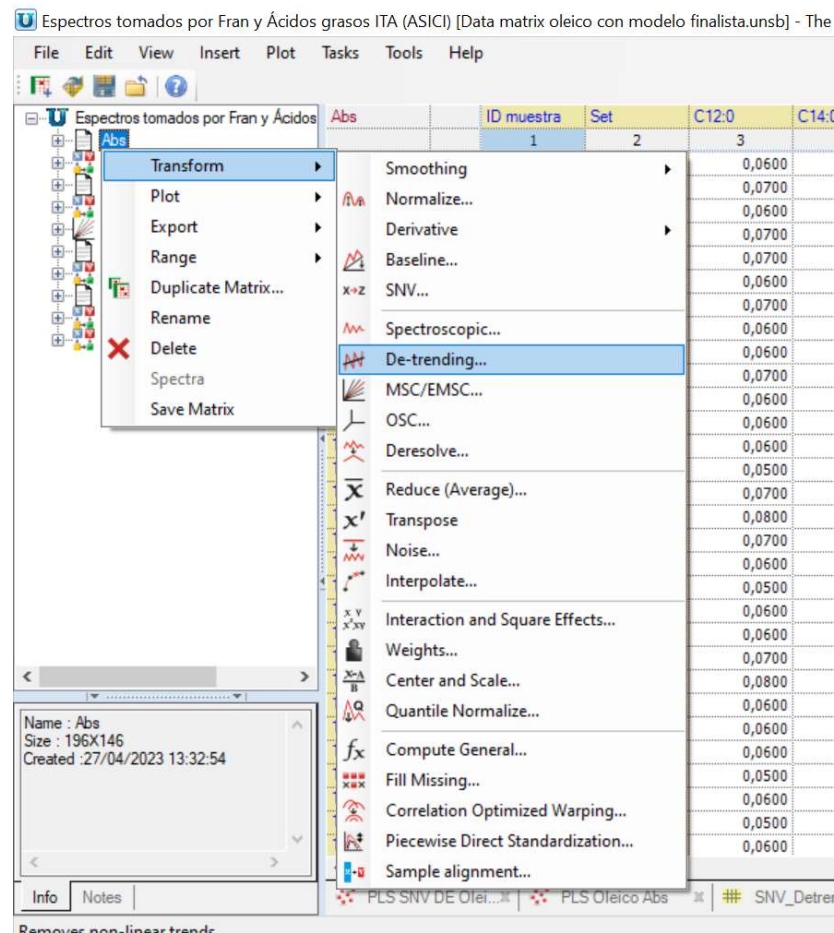
## Detrending (DT)

*Detrending (DT)* is a transformation which seeks to remove nonlinear trends in spectroscopic data, normally caused by scatter effect.



- Pre-treatments

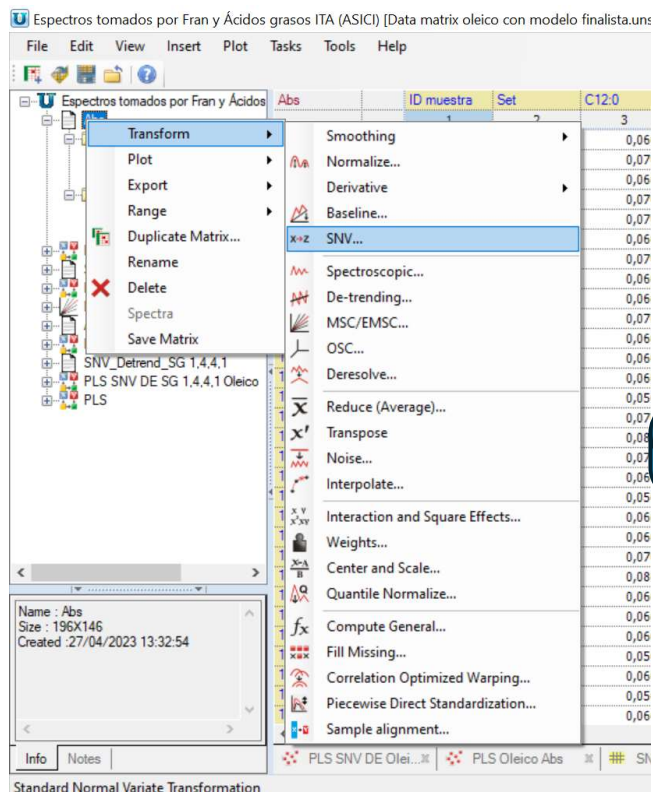
## Detrending (DT)



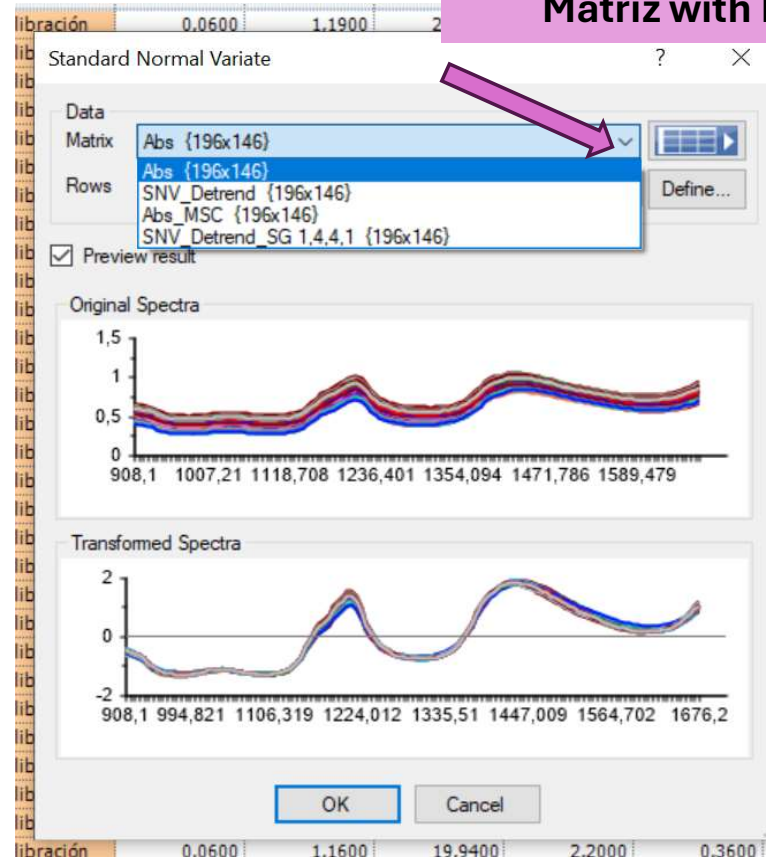
- Pre-treatments

## Standard Normal Variate (SNV)

It is usually applied together with the DT. The SNV is mainly used to correct baseline variations caused by scattering.



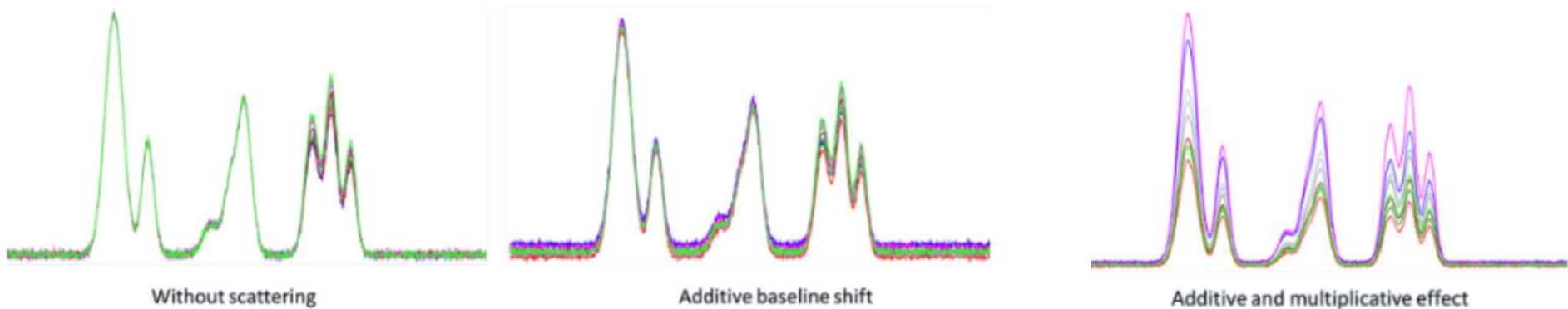
**Matriz with Detrending pre-treatment**



- Pre-treatments

## Multiplicative Scatter Correction (MSC)

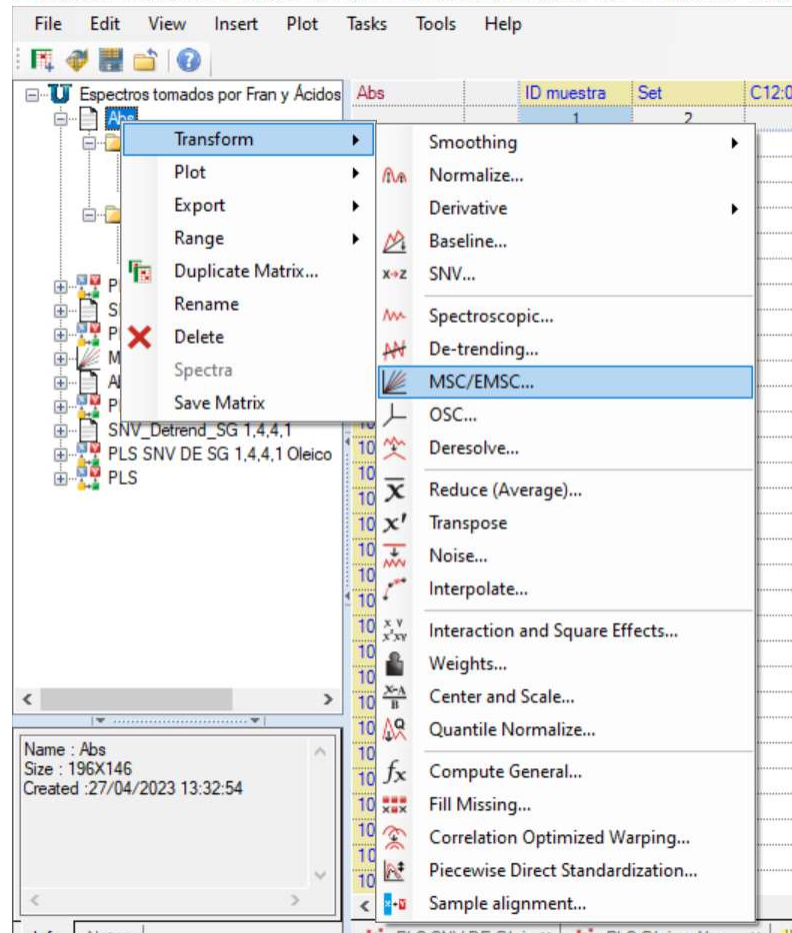
Multiplicative Scatter Correction (MSC) is a transformation method used **to compensate for pure additive and/or multiplicative effects** in spectral data.



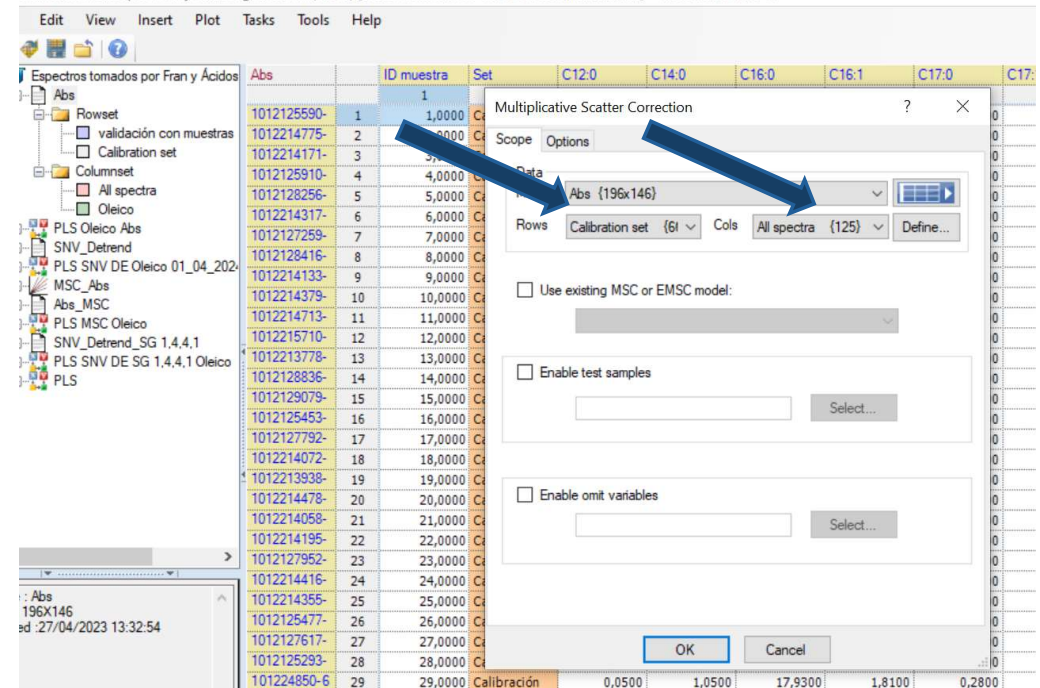
- Pre-treatments

## Multiplicative Scatter Correction (MSC)

Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Data matrix oleico con modelo finalist



Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Data matrix oleico con modelo finalista.unsb] - The Unscrambler X



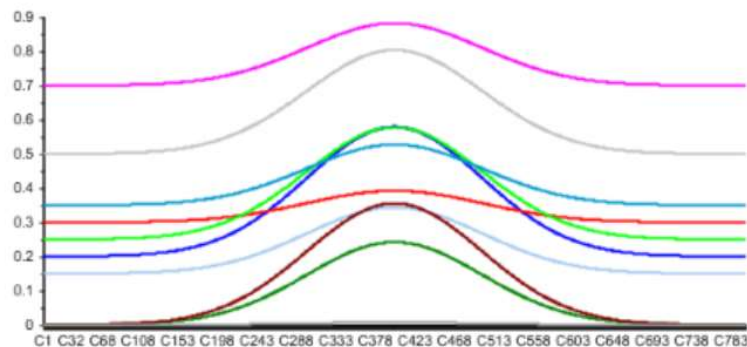


- Pre-treatments

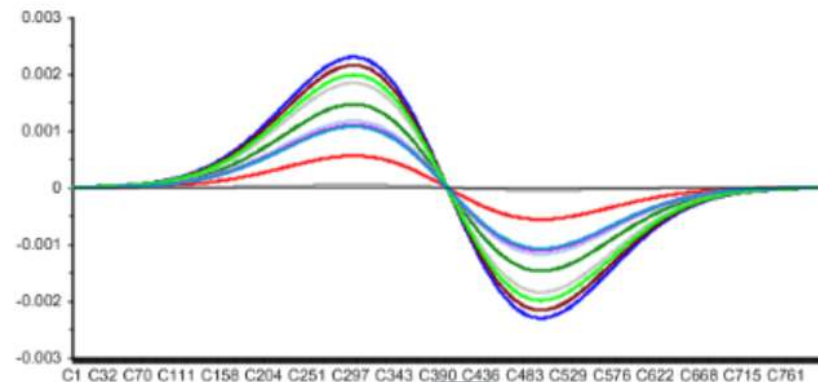
## Derivatives

Derivatives are applied **to correct for baseline effects (baseline shifts)** in spectra for the purpose of removing nonchemical effects. Derivatives may also aid in resolving overlapped bands which can provide a better understanding of the data, emphasizing small spectral variations not evident in the raw data.

**Raw data**



**First derivate**



The baseline offset has been removed under derivatization

The peak maxima in the raw data has now become a zero point in the derivative (The zero point can be explained by the fact that at a peak maxima (minima), the derivative is zero).

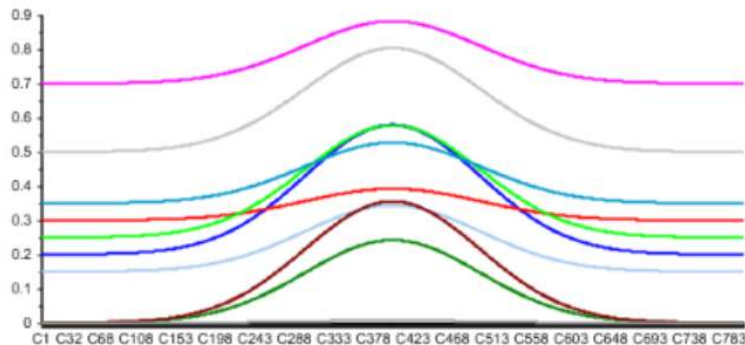
The **first derivative** eliminates **constant baseline shifts**

- Pre-treatments

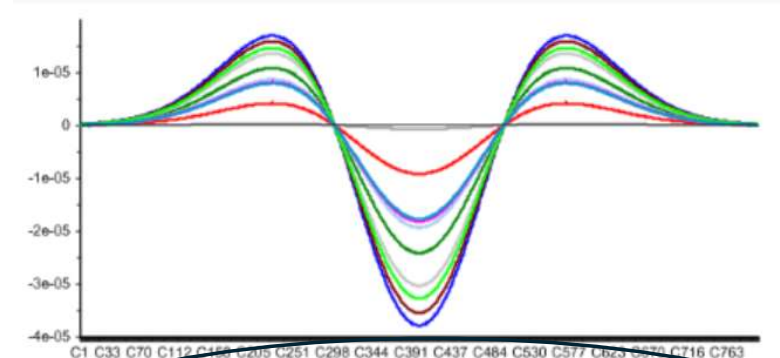
## Derivatives

Derivatives are applied **to correct for baseline effects (baseline shifts)** in spectra for the purpose of removing nonchemical effects. Derivatives may also aid in resolving overlapped bands which can provide a better understanding of the data, emphasizing small spectral variations not evident in the raw data.

**Raw data**



**Second derivate**



The second derivative is a measure of the change in the slope of the curve. In addition to removing pure additive offset, it is not affected by any linear “tilt” that may exist in the data.

**The second derivative can help resolve nearby peaks and sharpen spectral features**

The **second derivative** eliminates **shifts that vary linearly with wavelength**

- Pre-treatments

## Derivatives

### The Savitzky-Golay method

Espectros Uva NIRCROP 2023 [Data Matrix uva.unsb] - The Unscrambler X

**Derivative order**

**Polynomial order**

**Smothing points**

Derivative order: 1st Derivative

Polynomial order: 1

Number of left side points: 4

Number of right side points: 4

Number of smoothing points is 9

☒ Symmetric kernel

Transformed Spectra

OK Cancel

1,4,4,1

2,5,5,2

- Pre-treatments

## Remake the model on the pretreated matrix

Espectros tomados por Fran y Ácidos grasos ITA (ASICI) [Data matrix oleico con modelo finalista.unsb] - The Unscrambler X

File Edit View Insert Plot Tasks Tools Help

Espectros tomados por Fran y Ácidos

- Abs
- PLS Oleico Abs
- SNV\_Detrend
- PLS SNV DE Oleico 01\_04\_202
- MSC\_Abs
- Abs\_MSC
- PLS MSC Oleico
- SNV\_Detrend\_SG 1,4,4,1
- PLS SNV DE SG 1,4,4,1 Oleico
- PLS

SNV_Detren	ID muestra	Set	C12.0	C14.0	C16.0	C16.1	C17.0	C17.1
1012125590-	1	2						
1012214775-	2							
1012214171-	3							
1012125910-	4							
1012128256-	5							
1012214317-	6							
1012127259-	7							
1012128416-	8							
1012214133-	9							
1012214379-	10							
1012214713-	11							
1012215710-	12							
1012213778-	13							
1012128836-	14							
1012129079-	15							
1012125453-	16							
1012127792-	17							
1012214072-	18							
1012213938-	19							
1012214478-	20							
1012214058-	21							
1012214195-	22							
1012127952-	23							
1012214416-	24							
1012214355-	25							
1012125477-	26							
1012127617-	27							

Partial Least Squares Regression

Setup

- Model Inputs
- X Weights
- Y Weights
- Validation
- Algorithm

Model Inputs

Predictors

X SNV\_Detrend {196x146}

Rows

- Abs {196x146}
- SNV\_Detrend {196x146}
- Abs\_MSC {196x146}
- SNV\_Detrend\_SG 1,4,4,1 {196x146}

Responses

Y SNV\_Detrend {196x146}

Rows All Cols All Define...

Maximum components 7

☐ Identify outliers

☒ Mean center data

Cancel Back Next Finish

Name : SNV\_Detrend  
Size : 196 X 146  
Created : 01/04/2024 10:10:56  
Applied transformations : 2

- Pre-treatments

Pretreatments are only applied to **spectral variables**

Pretreatments are applied to the **calibration set**

The selection of the most appropriate one is purely experimental



Archivos de Zootecnia

Journal website: <https://www.uco.es/ucopress/az/index.php/az/>

Comparación de herramientas quimiométricas de clasificación para la identificación de grasa perirrenal en corderos

Agudo, B.<sup>1</sup>; Delgado, J.V.<sup>2</sup>; López, M.M.<sup>3</sup> y Rodríguez, P.L.<sup>4\*</sup>

<sup>1</sup>Ovino del Suroeste Soc. Coop. Ltda. (OVISO), Villanueva de la Serena, Badajoz, España.

<sup>2</sup>Departamento de Producción Animal, Facultad de Veterinaria, Universidad de Córdoba, España.

<sup>3</sup>Departamento de Producción Animal, CYCITEX, Finca La orden, Gobierno de Extremadura, España.

<sup>4</sup>Departamento de Producción Animal y Ciencia de los Alimentos, Facultad de Veterinaria de Cáceres, España.





## **Desarrollo de modelos predictivos mediante tecnología NIRS**

*Tecnologías avanzadas, innovadoras y digitales para el sector agroalimentario de la EUROACE*

*Évora, 18<sup>th</sup> and 19<sup>th</sup>, novembre, 2024*