

Introduction:

Near Infrared Reflectance spectroscopy is best performed in the 1900 to 2500nm region of the electromagnetic spectrum. Within this spectral region, Proteins, Amines, Amides (N-H), Moisture, Alcohol, Phenols (O-H) and Aliphatic Hydrocarbons, Aromatic Hydrocarbons Oils, Fats (C-H) absorb NIR energy. Using 0 - 45 degree illumination and detection optics, as shown in figure 1, provides a means of collecting NIR spectra from samples such as the Raw Materials used in Pharmaceutical manufacturing. Using Discriminant Analysis software provides a means to collect diffuse reflectance spectra from powdered samples and make a very accurate and precise means of identifying materials based on their NIR spectrum.



The MultiScan Series 4000 FTNIR Spectrometer offers several advanced features for performing Discriminant Analysis of Raw Materials: including superior wavelength accuracy and precision, variable resolution, faster scanning speed and immunity to stray light or external light.

The rotating sample dish provides a means of collecting at least 10 scans from a sample of powdered materials. This ensures a good average spectrum from the sample for comparison with the spectral library of known samples.

Typically the Discriminant Analysis takes less than 1 minute.

This application note illustrates the Discriminant Analysis of six materials using a library of materials commonly used in the Pharmaceutical industry.

Procedure:

Samples of raw materials were scanned using the Series 4000 FTNIR Spectrometer and a 10mm deep sample dish. The resolution for the scans was set at 64cm-1. 10 sub scans were collected for each sample and saved as the average spectrum for each sample into separate library files.



Figure 1 shows the spectra of the raw materials;

NTAS (Near Infrared Analysis Software) includes the Discriminant Analysis routine. Figure 2 shows the Discriminant Analysis window.

9 (Discriminant A	rialyste	- 8
File Options Scenning Windows Help			
Select the library files ADD/RPC ADD CSV EDTA CSV EDTA CSV FLOUG FACH CSV METHIONNE CSV OULARD CSV FOLYFROY CSV RICE FLOUR CSV RICE FLOUR CSV WHEY POWDER CSV GULLOSE CSV FROPUL PARAMETRICSV SODRUM ASCOPIBATE CSV	3	NO DATA MOREMBLE	
Add File Del File			
Results			
× Warelength 10000	Ourrent Scan:		Discriminate using
Y Wavelength 16017 🛓	• •		C 1st Derivative
Qlusters	Get Scan	Calibrate	Quit
Calibration Set Ready			4.08 PM 18/02/2015
	1 🖻 🧔 🚿 📰 🗉	. 🐺 🧏 🐨 🧿	📓 🧀 - 🛯 🖬 🖬 🕹 400 PM

Figure 2. Discriminant Analysis Window in NTAS.

Once the library files have been loaded, the Calibrate button is pressed. The software computes the Covariance Matrix for all the sample files. This is a complex multi dimensional vectorial space that generates a single resultant vector to represent each sample.

When an unknown sample is scanned, the Discriminant Analysis software computes the resultant vector for the new sample scan and then calculates the Mahalanobis Distance from this vector to every other vector in the multi dimensional vectorial space. The lower the Mahalanobis Distance from the unknown sample vector to the other vectors, the more likely that the two vectors are from the same sample.

The software then shows a list of 5 samples with an increasing Mahalanobis Distance between the library spectrum and the unknown spectrum. A perfect match should see a Mahalanobis Distance of 0, however the number is always a positive number.

Results:

Figure 3 through 9 show the results of performing the Discriminant Analysis routine on seven samples of materials which are included amongst the library files, ie, EDTA, Glucose, Sodium Ascorbate, Methionone, Poly Parabenzaldehyde and Whey Powder.

The Mahalanobis Distance for each analysis are listed below.

Selected Material	Mahalanobis Distance	Next Material	Mahalanobis Distance
EDTA	7.38	Poly Parabenzaldhyde	114479
Poly Parabenzaldhyde	6.99	Methionone	2296717
Glucose	26.34	Poly Parabenzaldhyde	2249331
Whey Powder	15.80	Poly Parabenzaldhyde	1170226
Methionone	6.03	Sodium Ascorbate	1465663
Citric Acid	19.51	Poly Propylene	707594

		CONTRACT Reality .		
Options Statistics Mildow	L Here			
lect the library files:	15			A
THONINE CEV LYPROP CEV	1.58			12
RTRATE CSV TRIC ACID CSV	8.64 -		2.42	
DUM ASCORDATE CRY	639-			
	131-		1	1
	128			
	820-			
	1.1.1			
Add File Del File	E17 -		N	
the second s	617 - 611 -	~		
HAUTS:				
isults:	ett - eos -	20 1000 0000 010 010 020 044 700 770 200 040 122 700 000	GEO 6400 CON CIO 1000 CON 5700 5100 1100 500 401 4400	HC 00 10
RUITS D'AL METHONNE CSV 10500T TARTRATE CSV 12 284 AGCOREC AGE CS	011 - 005 - 1000 0010 0020 9440 0	20 SOT MOD OTHER HTS COM 644 750 778 750 746 752 750 669 Current Science EDTA	Discriminate using	a desta and the second second
Halts 12 Year METHIONINE CSV 10 YOM ASCORDE AND CS 24 WW OTRICACO CSV	000 1000 001 000 044 0 10000 011 000 044 0	cto tone allos onte lette color et-le realo ritre rello rade rizze realo allos Current Scilen	Discriminate using	a desta and the second second





Figure 4. Poly Parabenzaldehyde

elect the library files:													
BCORBIC ADD CSV ITFRC ADD CSV	1.07												1
DTAICSV JOUR SHOM CEV ETHONINE CEV	076												1
LEACHLORY DLYPROP.COV	in.											A	
CERLOUR CSV INTRATE CSV											1	$\langle \rangle$	al
HEY POWDER/CBV UVCGBE CBV ROPYL PARABEN CBV	079						-				1	1	1
DOILIN ASCORBATE CEV	0.64						1	1			1	1	Ę
	0194						1				1		
	1000												
							1		-	-	/		
	2.0						1		-	1			
							1		1	1	J		
esults:	2.0						1		1	1]		
esults:	0.0 6.2%		~		~	لر	1		-	1)		
sults: 2.4555.0000 PROPYLENAPA MARKED SISSO PROPYLENAPA II ALTIS SISSO PROPYLENAPA	0.0- 6324 4.8- 0.15-		~			لے			~	~)		
sults: 2.4555.0000 PROPYLENAPA MARKED SISSO PROPYLENAPA II ALTIS SISSO PROPYLENAPA	0.0 6.2%		110 0450 000	0 (140 1990)	201 100 120	J	NALO GLAI NOO	0 6490 6290 9			1015010 5000 4		
594155 249255,0000 (HICOY) L NAVA 249255,0000 (HICOY) L NAVA 24955,0000 (HICOY) L NAVA 24950,0000 (HICO) R CACANI 24950,0000 (HICO) R CACANI	0.09 0.29 0.15 1000546 500 200 000	CLO SETO ET LA DETO AN CL	utrant Sc	an'	NOR 1000 1200		1010 CTN 1000	1 640 (28) 5			Discrimit	nate using	140 499 4130
Isalts: 2.4522.0000 PROPYL PAPA 2.45122.0000 PROPYL PAPA 4170110.0000 PROPYL PAPA	0.09 0.29 0.15 1000546 500 200 000	CLO SETO ET LA DETO AN CL	KF 8 9450 COB	an'	NOR 7880 750	0 7799 7220 7089		1000 CER 1			Discrime Absort	nate using bance	140 499 4130
594155 249255,0000 (HICOY) L NAVA 249255,0000 (HICOY) L NAVA 24955,0000 (HICOY) L NAVA 24950,0000 (HICO) R CACANI 24950,0000 (HICO) R CACANI	10000 1 10000 1 10000 1	CLO SETO ET LA DETO AN CL	utrant Sc	an') • 7 10 1 20 7 10					Discrimer * Absort 1 1 U De	nate using	140 499 4130
esuffs: 204522.0800 PHOP1 LINE 41 NET 0800 PHOP1 LINE 43 RET 0800 PHOP1 LINE 45 RET 0800 PHOP1 LINE X Wavelongth	10000 1 10000 1 10000 1	CLO SETO ET LA DETO AN CL	urrent Sci LUCOSI	san Eső							Discrimer * Absort 1 1 U De	nate using trate using trance erivative	140 4250 41

Figure 5. Glucose

Options Sciences Working of the library files: OPERCACE CEV INCACE CEV INCACE CEV INCACE CEV INCACE CEV INCACE CEV INCACE V VIENDECEV	0.45-		\mathcal{N}
ERLOUR CSV TRATE CSV EV POWDER CSV OOSE CSV PNL FARABER CSV IM ASCORBATE CSV	0.72-		N
	20.23		A 1
	2018		
	039- 005- 001		
ults: Totacian Price Levie 1012/2010/Price Levie 1015/2010/Price Levie 1015/2010/Price Levie	0.14 0.05 1000546 500 500 000 10	22 XY BIT IN AN AND AND AND AND AND AND AND AND AND	na nee ee ee Discriminate using
NIS ACTRONO PROPIDIAL TOTO CONTINUENTI PARA TOTO CONTINUENTI PARA TOTO CONTINUENTI TOTO CONTINUENTI ATTITIZZO CONTINUENTI ATTITIZI ATTITIZZO CONTINI ATTIT	014 039 04 0 1000946 400 400 000 0 1000946 400 400 000 0	convertent and was one must be real the real new new ten too the atto one and one con- Current Scient	nio tee ete ero teo toto toto son sen ate ete ate ete ate Discriminate using





Figure 8. Citric Acid

Discussion:

Discriminant Analysis is a powerful tool to perform library searches and to identify materials based on their Near Infrared spectra. The process of developing a library file is very simple. Scan 3-5 examples or batches of a material so that the spectra represent the day to day and batch to batch variations in the spectral scans. Store these into a unique library files. Over a period of time, add more spectra from different batches and different suppliers in order to build up a robust library of sample spectra. However a set of 30-100 library files can be developed in a few days using retained samples from batches of each raw material.

To analyse a sample using the Discriminant Analysis routine is as simple as filling the sample dish and pressing the Scan button. The software will collect the scans and display the results within 60 seconds.

The results shown in this study illustrate the very large differences in Mahalanobis Distances between different raw materials. It is therefore considered that variations in batches will not result in a false identification.

The Discriminant Analysis software can also be used to define a range of Mahalanobis Distances within which a specific raw material is considered Accept or Reject. For example, if the identification is correct but the Mahalanobis Distance is greater than a set value, ie, 2000, then the software can show a Red Button with Reject. If the Mahalanobis Distance is less than the set value, the software displays and Green Button and Accept.