



How to use UltraFine+[®] Next Gen Analytics QA/QC – Duplicates

Why is QA/QC important?

The data quality (i.e., precision and accuracy) of your soil analyses may impact how you interpret your results. Data quality interpretation can be a tedious manual process, especially when working with many hundreds or thousands of analyses. To assist you, we have developed a range of statistical QA/QC metrics to summarise the analytical quality of your duplicates in the UltraFine+[®] Next Gen Analytics workflow outputs. The outputs are designed to provide an overview of the repeatability of the sampling and analysis process. We use a “traffic light system” to classify measures of repeatability of the analyses to give you a rapid, first-pass indication of how precise your analyses are.

Where do I find the QA/QC for my duplicates?

The file *duplicates_QAQC.xlsx* is in your *Data Package* under *1 Data* → *QAQC*. The file contains four tabs for geochemistry, sizing, VNIR and FTIR analyses. Automation of the QA/QC workflow requires availability of UltraFine+[®] duplicate analyses and consistent formatting of input data. Where this is not provided, your *Data Package* will not include QA/QC on duplicates.

The “Traffic Light” System

The traffic light system for duplicate pairs is intended to give you a rapid overview of the repeatability of your analyses and highlight anything that you may want to investigate further. For your duplicate analyses, the traffic light system has four (geochemistry and sizing) or three (VNIR and FTIR) colours: If everything is green, there is nothing to worry about. The colours yellow, (orange) and red indicate that there may be a problem on an escalating scale, respectively. It is important to note that analytical precision close to the detection limit is typically compromised. We therefore indicate values that are less than 10 times the detection limit in *purple* font. Data highlighted in *green* font indicates values below the detection limit and these have been replaced with half the detection limit.

Geochemistry

If all of the cells in your duplicate pair summary (Figure 1) are green, there is nothing to worry about. If you see other colours in your duplicate pair summary, you can access more information on individual duplicate pairs by clicking on the + symbol to open the drop-down menu (Figure 1). If you see consistent red flags for a specific duplicate pair, especially where there are also yellow and orange flags this is a strong indication that your duplicate pair may in fact not be a duplicate pair. A reasonable first step would be to confirm your meta data (sample ID’s, coordinates and field notes). If specific element analyses are consistently flagged across multiple duplicate pairs, and you are satisfied that your duplicate meta data is correct, you may want to contact the laboratory for clarification on the precision of the analyses.

Figure 1: Example duplicate pair QA/QC summary output and expanded drop-down menu for individual duplicate pair geochemical analyses.

What are the different colours for my geochemistry duplicates based on?

We developed our traffic light system by conditionally formatting the half absolute relative difference for duplicate pair geochemical elements and EC, and on the absolute difference for duplicate pair pH values. The specific rules for colour coding are outlined in Table 1 and take the reliability of the UltraFine+[®] analytical method into account.

Table 1: Conditional formatting rules for geochemical duplicate pairs analysed with UltraFine+[®]. *Ca, Ce, Fe, S, Ta, Br, Sc; **Cr, Ti, Zr, W, Te, Sn, Se, La, Hf; ***Au, Pt, Pd; ****Ag, Al, As, Ba, Be, Bi, Cd, Co, Cs, Cu, Ga, Ge, In, K, Li, Mg, Mn, Mo, Nb, Ni, Pb, Rb, Re, Sb, Sr, Th, Tl, U, V, Y, Zn; †Field will be grey if analyses are <300 µS/cm

Problematic elements*	Resistate elements**	Precious metals***		Hg		EC	pH	All other elements****
		If >10 ppb	If <10 ppb	If >0.2 ppm	If <0.2 ppm	>300†	(unit difference)	
<15 %	<50 %	<10 %		<10 %		<50 %	<0.5	<10 %
>15 %	>50 %	>10 %	>10 %	>10 %	>10 %	>50 %	>0.5	>10 %
>30 %	>75 %	>15 %		>15 %		>75 %	>0.75	>15 %
>45 %	>100 %	>20 %		>20 %		>100 %	>1	>20 %

Sizing

Similar to your geochemistry data, if all cells in your duplicate pair summary are green, there is nothing to worry about. If all three sizing parameters are red, or red and orange, and you are satisfied that your duplicate metadata is correct, you may want to contact the lab for clarification on the precision of the analyses. You can view individual duplicate pairs by clicking on the + symbol to open the drop-down menu (Figure 2) and refer to the table at the end of this document for an explanation of the calculated parameters.

Figure 2: Example QA/QC summary output for duplicate pair sizing analyses. You can view individual duplicate pair comparisons by clicking on the + sign.

What are the different colours for my sizing duplicates based on?

Routine protocols for QA/QC of sizing data at the post-processing stage are not universally established. Our team has reviewed duplicate analyses throughout the Next Gen Analytics project to derive conditional formatting rules for the sizing parameters in the automated QA/QC outlined in Table 2. Duplicate pair QA/QC for sizing data has been developed on the 10th, 50th and 90th percentile of the particle size.

Table 2: Conditional formatting rules for duplicate pair sizing analyses.

d_0.1_µm	d_0.5_µm	d_0.9_µm
<10 %	<10 %	<10 %
>10 %	>10 %	>10 %
>15 %	>15 %	>15 %
>20 %	>20 %	>20 %

VNIR

Similar to the geochemical analyses, if all cells in your duplicate pair summary are green, there is nothing to worry about. If some scalars are flagged as yellow or red you can access more detailed information on individual duplicate pairs by clicking on the + symbol to open the drop-down calculations (Figure 3). Red flags for individual scalars across multiple duplicate pairs can indicate that this scalar may not be reliable, this should be considered when you interpret your analyses. These values or anomalous results should be used with caution and checked against other parameters if you are looking to use these numbers to, for example, derive a new target. If you see consistent red flags for multiple scalars of a duplicate pair, especially where there are also yellow and orange flags, and you are satisfied that your duplicate pair metadata is correct, you may want to contact the laboratory for further clarification.

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
30	+	SampleID	Job No	Client Nar	Date Repc	Duplicate	ClientID	Client Ref	MinGrp1	MinGrp2	FeOx	Hem_Geo	Kln_abun	Kln_cryst	WM_AS_abun	WM_AS_comp	Fe_Kln	Chi	FeMgClay	OH_mafic	Paly	Colour	Hue	Saturator	Intensity	NIR_TSG_ver
31	IsEqual	mode							Same	Same																Same
32	HARD%	min									0.47		0.34	0.05			4							0	0.35	
33	HARD%	mean									1.18		2.44	0.08			6.17							0.34	1.82	
34	HARD%	max									2.86		5.85	0.14			8.33							0.68	4.88	
35	AbsDiff	min										0														
36	AbsDiff	mean										0.55														
37	AbsDiff	max										1.4														
38	Angle	min																					0.2			
39	Angle	mean																					0.27			
40	Angle	max																					0.33			

Figure 3: Example duplicate pair QA/QC summary output for duplicate pair NIR analyses. Click on the + sign to view individual duplicate pair comparisons.

What are the different colours for my VNIR duplicates based on?

Routine protocols for QA/QC of VNIR data at the post-processing stage are not universally established. Our team has reviewed duplicate analyses throughout the Next Gen Analytics project to derive conditional formatting rules for all scalars as outlined in Table 3. Where no data is reported for specific scalars in any duplicate pairs (i.e., these parameters were not detected in the specific samples) cells will be blank. In the example above, this is the case for the parameters WM_AS_abun, WM_AS_comp, Chi, FeMgClay, OH-mafic and Paly. Hence, there is no data in columns N, O, Q-T. Column U is also empty, since no comparison is applied to Colour, because the Munsell Colour is derived from Hue, Saturation and Intensity.

Table 3: Conditional formatting rules for VNIR duplicate pairs. *Fe-Kln, Chi, FeMgClay, OH-mafic, Paly

MinGrp1	MinGrp2	FeOx	Hem_Geo	Kln_abun	Kln_cryst	WM_AS_abun	WM_AS_comp	Other scalars*	Hue	Saturation	Intensity
Same	Same	<8%	<3nm	<7%	<0.35%	<6%	<1 nm	<5%	<2%	<2%	<7%
N/A	Different	8-16%	3-4.5nm	7-13%	0.35-0.7%	6-12%	1-1.5 nm	5-10%	2-4%	2-4%	7-14%
Different	N/A	>16%	>4.5nm	>13%	>0.7%	<12%	>1.5nm	>10%	>4%	>4%	>14%

FTIR

Similar to the geochemical analyses, if all cells in your duplicate pair summary are green, there is nothing to worry about. If some scalars are flagged in red in your duplicate pair summary, you can access more detailed information by clicking on the + symbol to open the drop-down menu to access more information on individual duplicate pairs (Figure 4). If you see consistent red flags for a duplicate pair, especially where there are also yellow and orange flags, and you are satisfied that your duplicate metadata is correct, you may want to contact the laboratory for clarification on the precision of the analyses. Red flags for individual parameters across multiple duplicate pairs can indicate that this parameter may not be reliable. This should be considered when you interpret your analyses. If there are consistent red flags across multiple duplicate pairs for multiple scalars, you may want to contact the laboratory for further clarification.

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M
26	+	SampleID	Job No	Client Nar	Date Repc	Duplicate	ClientID	Client Ref	Clay_wt%	Qtz_wt%	Carb_wt%	TOC_wt%	Gibbs_ind	FTIR_TSG_ver
27	HARD%	min							0.06	3.55				
28	HARD%	mean							2.85	5.85				
29	HARD%	max							7.37	7.69				
30	RankMetr	min											0	
31	RankMetr	mean											0	
32	RankMetr	max											0	
33	IsEqual	mode												Same

Figure 4: Example QA/QC summary output for duplicate pair FTIR analyses. Click on the + sign to view individual duplicate pair comparisons.

What are the different colours for my FTIR duplicates based on?

Routine protocols for QA/QC of FTIR data at the post-processing stage are not universally established. Our team has reviewed duplicate analyses throughout the Next Gen Analytics project to derive conditional formatting rules for all parameters as outlined in Table 4. Where values are below the detection limit for specific parameters in any duplicate pairs, cells will be blank (e.g., for carbonate and TOC in the Figure 4 example, hence no data is visible in the summary for columns J and K).

Table 4: Conditional formatting rules for FTIR duplicate pairs.

Clay %, Quartz %, Carbonate %, TOC %		Gibbsite Index
if values <5 %	if values >5 %	
<10 %	<20 %	Same
>10 %	>20 %	N/A
>20 %	>50 %	Different

What do the parameters for my duplicate pair calculations mean and how were they calculated?

Parameter	Formulae	Definition
HARD%	$100 \frac{ x_a - x_b }{x_a + x_b}$	Half relative difference between the respective value for duplicate pairs
HARD% mean		The mean of all half relative differences between the respective values for all duplicate pairs
RankMetric mean		Rank differences between two ordinal values, used here to compare the modal Gibbsite index (with ordinal values [NULL, 1, 2, 3])
IsEqual mode		Modal value of the equality check for all duplicates, indicating if the majority of duplicate pairs give equal or unequal values
AbsDiff	$ x_a - x_b $	Absolute difference between duplicate pair value
AbsDiff mean		The mean of absolute difference of a respective value for all duplicate pairs
Angle mean	$\cos^{-1} \left(\frac{\vec{u} \cdot \vec{v}}{\ \vec{u}\ \cdot \ \vec{v}\ } \right)$	Angular difference between two vectors; here used for the hue component of colour as encoded as an angle
<parameter> min		The minimum value of all respective parameter differences between the respective values for all duplicate pairs
<parameter> max		The maximum value of all respective parameter differences between the respective values for all duplicate pairs