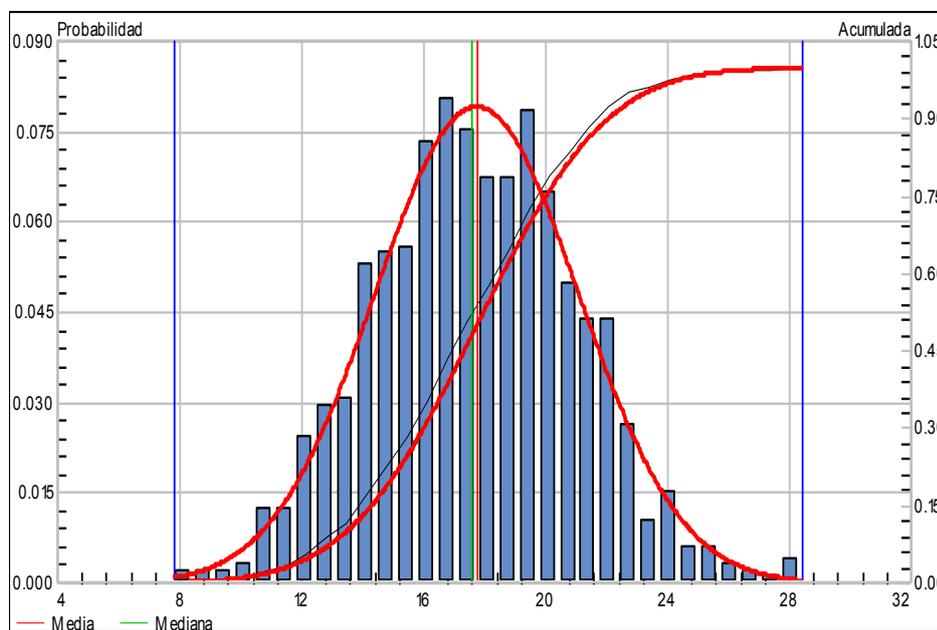


# HERRAMIENTAS ESTADÍSTICAS

Una simulación de Montecarlo permite remplazar un enfoque determinístico por uno estocástico introduciendo el grado de incertidumbre necesario al problema que se pretende resolver.

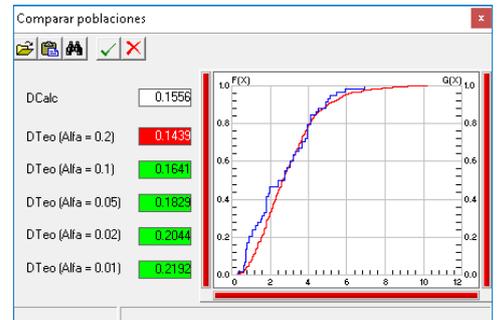
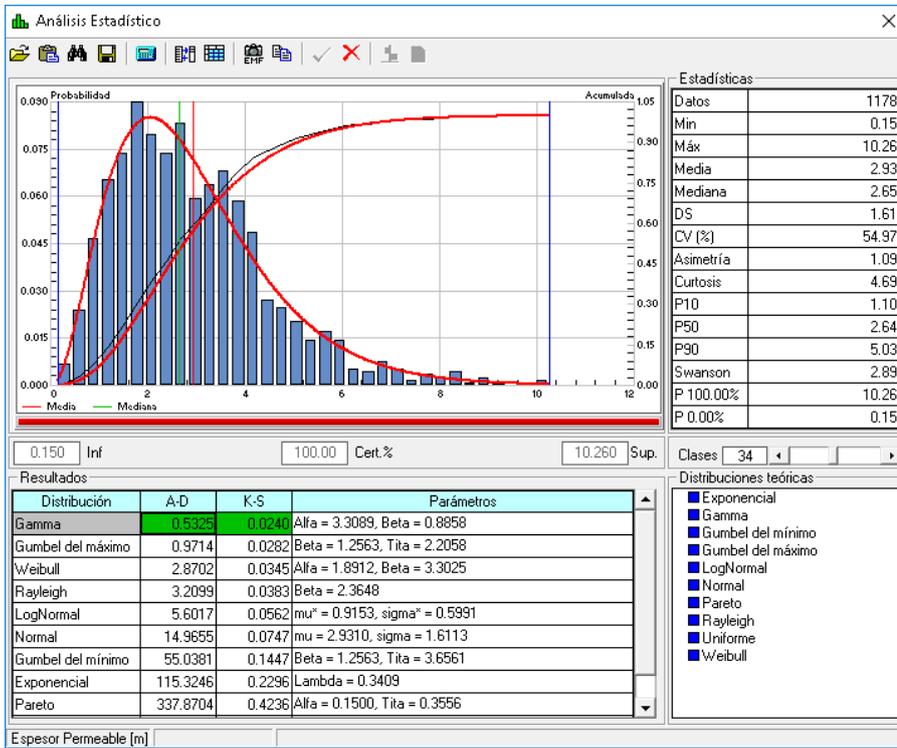
Sahara posee una amplia lista de conocidas distribuciones de probabilidad continuas y discretas que pueden ser utilizadas para definir las hipótesis de las variables involucradas en una simulación de Montecarlo. El conocimiento de las características de las funciones de densidad de probabilidad, el papel que juegan sus parámetros en sus cambios de forma y escala, y sus aplicaciones más frecuentes son fundamentales para garantizar la calidad de los resultados. Para ello, Sahara brinda la posibilidad de realizar un análisis estadístico sobre los valores de múltiples variables.

El módulo de Análisis estadístico permite graficar la distribución de los datos de alguna variable en particular y compararla con las distintas distribuciones teóricas disponibles. Para la distribución que mejor ajuste se pueden obtener los parámetros característicos y guardarlos para ser consultados posteriormente o utilizarlos como hipótesis en una simulación de Montecarlo.



**Se pueden realizar análisis estadísticos y simulaciones de Montecarlo utilizando datos dentro de un proyecto Sahara o simplemente utilizando datos externos.**

**Sahara**



**Comparar poblaciones.** Se puede ver la comparación de dos muestras. En este ejemplo no hay evidencia suficiente para inferir que las dos muestras derivan de diferentes poblaciones para los niveles de confianza correspondientes a  $\alpha = 0.10, 0.05, 0.02$  y  $0.01$ , pero sí para  $\alpha = 0.20$ .

**Análisis estadístico.** Ejemplo en el que se estudia el espesor permeable de una capa en particular para un grupo de pozos. Se puede ver que se han seleccionado todas las distribuciones teóricas para realizar el estudio y entre todas ellas, la que mejor representa los datos, según los test de bondad de ajuste, es una distribución Gamma con  $\alpha=3.30$  y  $\beta=0.89$ . Se observa sobre el histograma la representación de la distribución Gamma y su acumulada.

## Con el módulo de Análisis estadístico se puede estudiar la distribución de datos de variables dentro de un proyecto Sahara o también realizar el análisis para datos externos a un proyecto.

La ventana de análisis estadístico permite realizar análisis para variables dentro de un proyecto de Sahara o para datos externos. Adicionalmente, la misma puede utilizarse para analizar los resultados obtenidos de una simulación de Montecarlo.

En esta ventana se pueden visualizar datos estadísticos asociados a la muestra, como los valores máximo, mínimo, media, mediana, la desviación estándar de los mismos y los valores asociados a los percentiles 10, 50 y 90. El histograma puede modificarse utilizando la cantidad de clases que se crean necesarias para una óptima visualización. Además, ajustando dos barras verticales podrán definirse valores límites para evaluar el porcentaje de certeza de que un valor dentro de la muestra caiga dentro de los límites establecidos.

En la ventana se dispone de un listado con las distribuciones teóricas continuas y discretas disponibles. Las mismas pueden ser seleccionadas con el objetivo de intentar ajustar los datos de la muestra y encontrar aquella distribución y parámetros que logren el mejor ajuste. Para estimar los parámetros muestrales para cada una de ellas, Sahara utiliza los métodos de estimación por máxima verosimilitud y el método de los momentos, según el caso. Luego, para chequear la suposición de que un conjunto de datos pertenece a una determinada distribución de

probabilidades conocida, se utilizan dos test de bondad de ajuste: Anderson-Darling y Kolmogorov-Smirnov. Los mismos proponen una hipótesis nula, con cierto nivel de significación, que se acepta o se rechaza. Si la misma se rechaza, entonces se prueba que los datos no se distribuyen según la distribución propuesta. Si, por el contrario, se acepta, se establece que el conjunto de datos es consistente con la distribución propuesta y se mide el grado de concordancia entre la distribución del conjunto de datos y una distribución teórica determinada.

Luego de realizar el cálculo, las distribuciones seleccionadas se ordenan según el grado de ajuste, mostrando en color verde los parámetros de ajuste que representan el mejor de ellos. Estos resultados se pueden guardar para ser estudiados posteriormente o para ser utilizados como hipótesis en una simulación de Montecarlo.

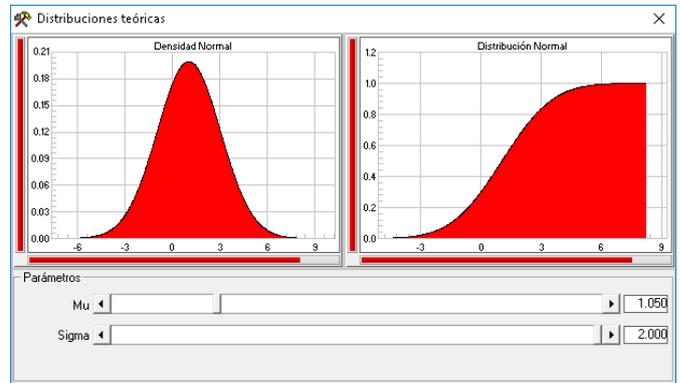
En este módulo también se permite realizar una comparación de poblaciones. Mediante este procedimiento es posible discernir, con criterio estadístico si dos muestras independientes pertenecen o no a la misma población. Este hecho es muy importante, por ejemplo, para asegurar la calidad de los datos con los que se armarán las series de datos históricos de una variable a utilizar en una simulación de Montecarlo. De este modo, se podrán obtener resultados más realistas. El método utilizado por Sahara para realizar este procedimiento es la prueba de Kolmogorov-Smirnov para dos muestras. La misma se basa en la comparación de las funciones de distribución acumuladas muestrales, y no se requiere supuesto alguno sobre las distribuciones de los datos comparados.

El objetivo de una simulación de Montecarlo consiste en obtener un resultado estocástico para cualquier problema que se plantea. Para lograr esto, los datos de entrada del modelo deberán ser distribuciones de probabilidad en lugar de números determinados. La asignación de la distribución a cada variable es uno de los pasos más complejos e importantes del proceso, ya que si las hipótesis están erradas nos llevarán a resultados alejados de la realidad.

Cuando deben determinarse las distribuciones para utilizar como hipótesis de cada variable, se encuentra disponible la opción de conocer las características de cada distribución teórica y cómo impacta la variación de sus parámetros. Esta funcionalidad es de gran ayuda a la hora de definir las hipótesis. Para cada una de las distribuciones continuas y discretas disponibles en Sahara, la ventana Distribuciones Teóricas provee:

- Gráfico de la función de densidad y distribución acumulada de la misma.
- Deslizadores para variar los parámetros que definen la distribución. Esta opción tiene como objetivo visualizar cualitativamente cuál es el efecto de los mismos sobre la forma y escalado de las funciones de densidad y distribución acumulada. Los intervalos de valores de los parámetros están acotados arbitrariamente entre valores predeterminados.

La técnica de Montecarlo consiste en plantear un modelo o ecuación, asignar a cada variable de entrada la correspondiente distribución de probabilidad y realizar el cálculo múltiples veces. Como parámetro adicional, se puede utilizar una matriz de correlación entre las variables. De esta forma, si existe relación entre algunas de ellas, la misma podrá tenerse en



**Parámetros de distribuciones teóricas.** En la figura se observa una distribución Normal con parámetros  $\mu = 1.050$  y  $\sigma = 2.000$ .

cuenta a la hora de realizar el cálculo. En cada corrida, para cada una de las variables de entrada se generan de manera aleatoria valores según la distribución de probabilidad asignada. Para asegurar la reproducibilidad de los resultados, se recomienda realizar un número grande de corridas. En cuestión de segundos, Sahara realizará los cálculos del orden de miles de corridas para arrojar un resultado en la ventana de Análisis estadístico. Para los resultados se pueden observar datos estadísticos y comparar la distribución del mismo con las distribuciones teóricas disponibles dentro de Sahara.

Existe la posibilidad de elegir la opción Hipercubo latino para realizar el muestreo de datos. Este método de muestreo estratificado permite disminuir el número de corridas a realizar, ya que segmenta la distribución de probabilidad acumulada y genera números aleatorios dentro de cada segmento, sin repetir el muestreo en ninguno de ellos.

Adicionalmente, se puede visualizar un gráfico de sensibilidad de tipo tornado, en el que se puede evidenciar el peso de cada una de las variables involucradas sobre el resultado final.

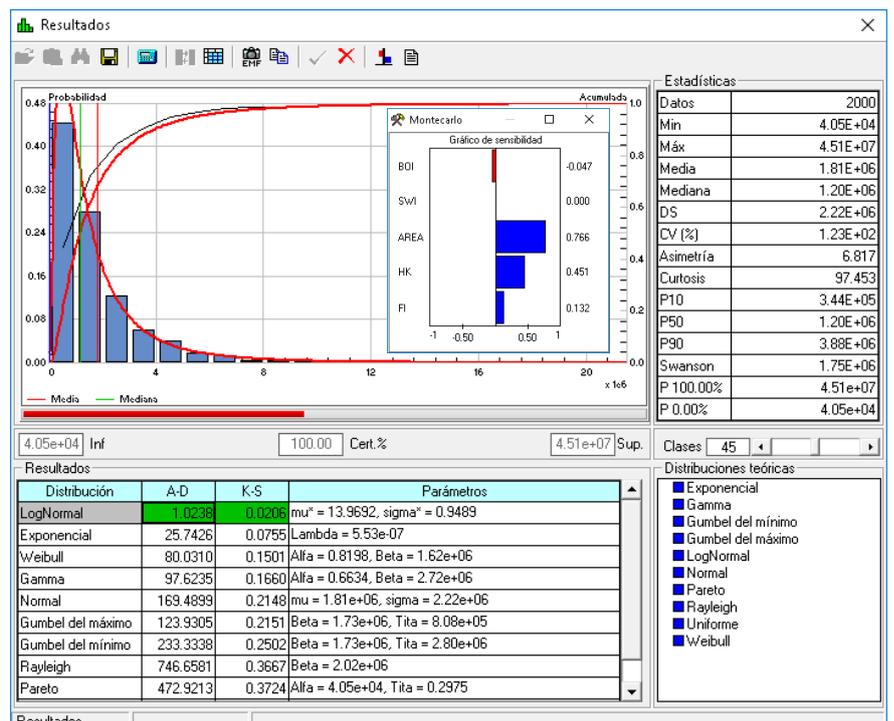
Var.	Distribución	Parámetros
FI	Normal	Mu = 0.1773, Sigma = 0.0335
HK	Gamma	Alfa = 3.3057, Beta = 0.8866
AREA	LogNormal	P10 = 1540000, P90 = 9900000
SWI	Triangular	P2 = 0.19, Más probable = 0.22, P98 = 0.25
BDI	Uniforme	Mínimo = 1.05, Máximo = 1.25

	FI	HK	AREA	SWI	BDI
FI	1	0.5	0.000	0.25	0.000
HK		1	0.000	0.000	0.000
AREA			1	0.000	0.000
SWI				1	0.000
BDI					1

**Montecarlo.** Definición de hipótesis para las variables de una simulación Montecarlo con el objetivo de analizar el volumen de hidrocarburo in situ. Además se visualiza la matriz de correlación de las variables.

**Resultados.** Se observan los resultados de la simulación. De todas las distribuciones disponibles, según los test de bondad de ajuste, el volumen de OOIP tiene una distribución Gamma con  $\alpha=2.88$  y  $\beta=12.30$ . Se observa sobre el histograma la representación de la distribución Gamma y su acumulada. Adicionalmente, se observa un gráfico tornado que indica que las variables con mayor peso sobre el resultado son el espesor y la porosidad.

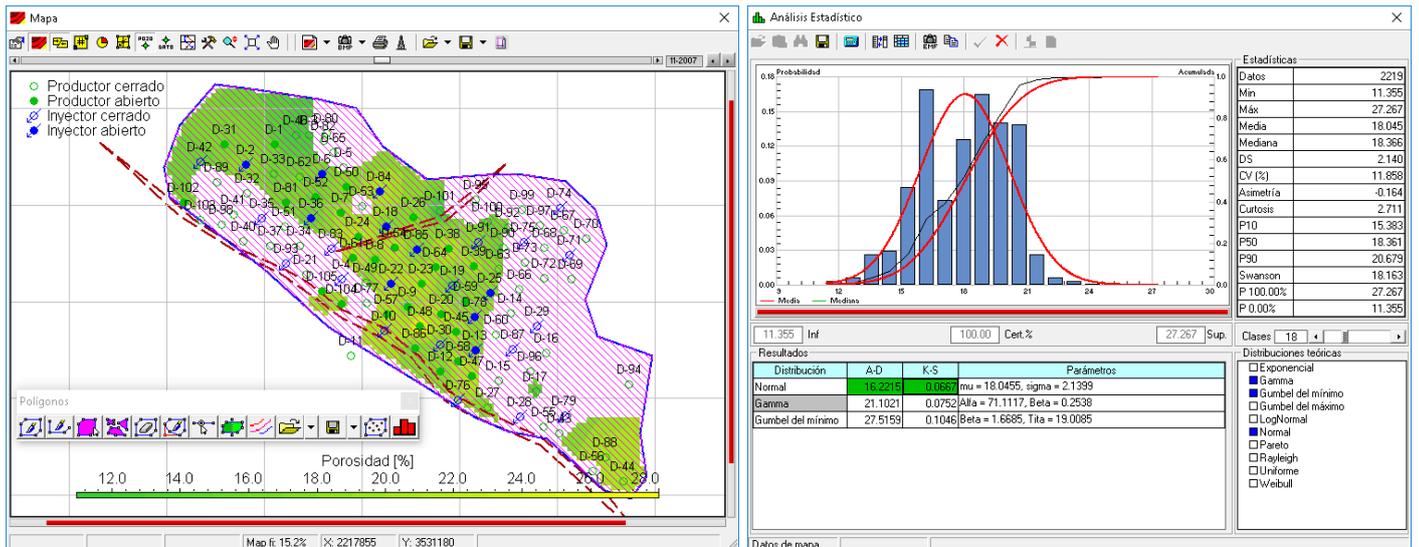


Distribución	A-D	K-S	Parámetros
LogNormal	1.0238	0.0206	$\mu^* = 13.9692$ , $\sigma^* = 0.9489$
Exponencial	25.7426	0.0755	Lambda = 5.53e-07
Weibull	80.0310	0.1501	Alfa = 0.8198, Beta = 1.62e+06
Gamma	97.6235	0.1660	Alfa = 0.6634, Beta = 2.72e+06
Normal	169.4899	0.2148	$\mu = 1.81e+06$ , $\sigma = 2.22e+06$
Gumbel del máximo	123.9305	0.2151	Beta = 1.73e+06, Tita = 8.08e+05
Gumbel del mínimo	233.3338	0.2502	Beta = 1.73e+06, Tita = 2.90e+06
Rayleigh	746.6581	0.3667	Beta = 2.02e+06
Pareto	472.9213	0.3724	Alfa = 4.05e+04, Tita = 0.2975

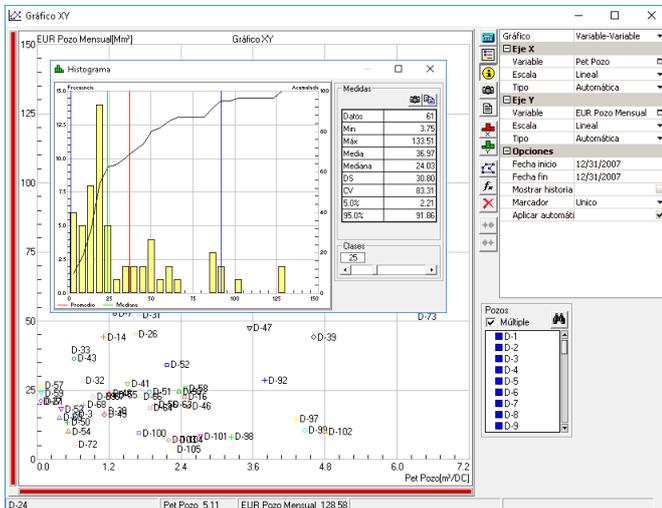
Estadísticas	
Datos	2000
Min	4.05E+04
Máx	4.51E+07
Media	1.81E+06
Mediana	1.20E+06
DS	2.22E+06
CV (%)	1.23E+02
Asimetría	6.817
Curtosis	97.453
P10	3.44E+05
P50	1.20E+06
P90	3.88E+06
Swanson	1.75E+06
P 100.00%	4.51E+07
P 0.00%	4.05E+04

- Distribuciones teóricas
- Exponencial
  - Gamma
  - Gumbel del mínimo
  - Gumbel del máximo
  - LogNormal
  - Normal
  - Pareto
  - Rayleigh
  - Uniforme
  - Weibull

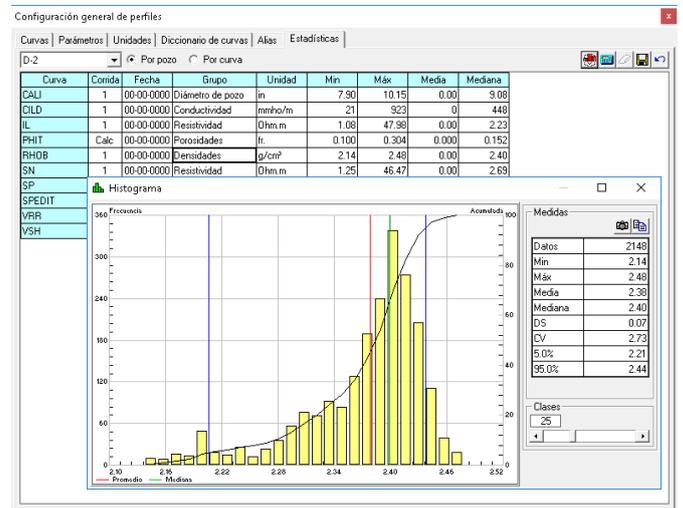
Una prestación adicional de la ventana de análisis estadístico es que la misma puede ser llamada desde muchas otras herramientas de Sahara, permitiendo realizar análisis para los distintos tipos de información:



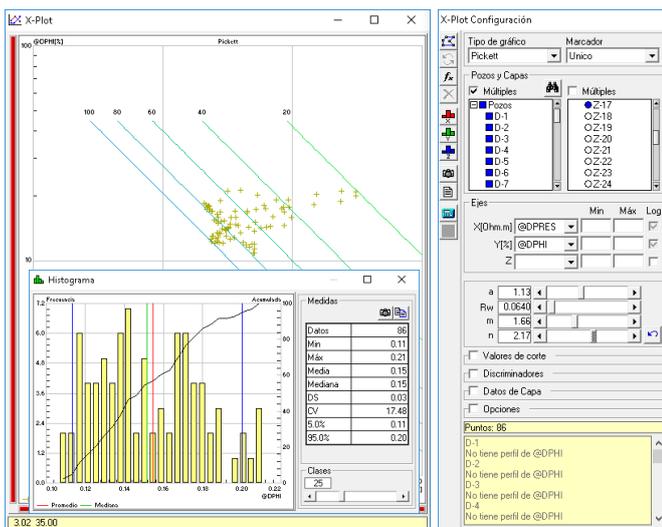
**Mapas:** realizando un polígono en la ventana mapa, se puede analizar el histograma de los valores de las celdas del mapa que se está visualizando.



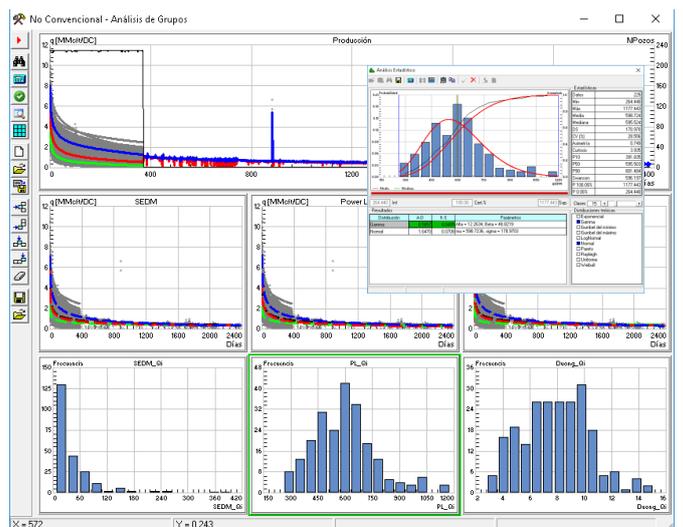
**Gráficos XY:** puede visualizarse el histograma correspondiente a los valores graficados en cualquiera de los ejes de los gráficos de la ventana.



**Configuración de Perfiles:** para cada uno de los perfiles de los pozos, se puede observar el histograma correspondiente y los parámetros estadísticos de la distribución. Permite detectar puntos anómalos.



**Cross-Plot:** puede visualizarse el histograma correspondiente a los valores de los perfiles graficados en cualquiera de los ejes de los gráficos de la ventana.



**No convencional - Analisis de grupos:** dentro de este módulo existen gráficas específicas para realizar análisis estadísticos. Desde cada una de ellas se puede acceder a la ventana de análisis estadístico.

