

Análisis de un Modelo Estadístico de Regresión Aplicado al Efluente de una Planta de Tratamiento de Aguas

CONTRERAS, Jannette*†, CRUZ, Braulio y PEÓN, Ricardo

Universidad Autónoma de Yucatán.

Recibido 3 de Julio, 2015; Aceptado 1 de Septiembre, 2015

Resumen

Los estudios de los tratamientos de tipo fisicoquímico de las aguas residuales son de gran importancia pues permiten a través de los datos obtenidos en laboratorio analizar modelos estadísticos que ayudan a describir el comportamiento del proceso así como identificar las variables que tienen mayor influencia sobre otras, en el presente estudio se analizaron los clarificados provenientes de diez pruebas de jarras tomando como muestra agua residual del efluente de una planta de tratamiento de aguas de tipo fisicoquímico; mediante un análisis estadístico de regresión se llegó a la conclusión de que la turbidez es una variable de la cual depende la demanda química de oxígeno de las mismas aguas.

Prueba de jarras, aguas residuales, regresión lineal, turbidez, SST.

Abstract

Studies of wastewater physicochemical treatments are important because through the data obtained in the laboratory they allow to analyze statistical models that help to describe the process behavior and identify the variables that have the greater influence on others, in this study clarifiers were analyzed from ten test jars using wastewater effluent from a sewage treatment physicochemical plant as sample; by means of a statistical regression analysis it is concluded that chemical oxygen demand depends on the turbidity variable of the same water.

Jar test, waste water, lineal regression, turbidity, SST.

Citación: CONTRERAS, Jannette, CRUZ, Braulio y PEÓN, Ricardo. Análisis de un Modelo Estadístico de Regresión Aplicado al Efluente de una Planta de Tratamiento de Aguas. Revista de Análisis Cuantitativo y Estadístico 2015,2-4: 271-277

* Correspondencia al Autor (Correo Electrónico: jannette.contreras@uady.mx)

† Investigador contribuyendo como primer autor

Introducción

El presente estudio estadístico, se basa en datos provenientes de análisis de laboratorio realizados a aguas residuales que se generan en una institución educativa; tales efluentes son recolectados de las aguas residuales de los diversos servicios que el campus ofrece, como son: cafeterías, baños, laboratorios y cualquier otro tipo de servicio que requiere del uso del agua. El agua residual que se genera de los sitios antes mencionados es tratada por medio de un proceso de tipo fisicoquímico el cual involucra la utilización de sustancias coagulantes que llevan a cabo la remoción de los contaminantes; la gama de estas sustancias es muy variada y cada coagulante posee diferentes propiedades fisicoquímicas haciendo que cada uno tenga diferente eficiencia en la remoción de los contaminantes.

Al momento del estudio la institución contaba con una planta de tratamiento de aguas residuales de tipo fisicoquímico, la cual debido al incremento de usuarios de las instalaciones del campus tales como servicios sanitarios, servicios alimentarios, centros deportivos, etc. recibe una mayor fluctuación de aguas residuales que a su vez contienen una mayor carga orgánica y de sólidos. En el año de 1986, la institución contaba con 10, 373 alumnos y 866 profesores; para el año de 1990, la cantidad de alumnos se incrementó a 16, 277 y de profesores a 1, 177. Para el año de 2003 la cantidad de alumnos llegó a la cifra de 16, 371 alumnos y la cifra de profesores a 1394.

En general se generan en promedio 775.76 m³ de agua residual al día, que son tratadas por la propia institución mediante un proceso de coagulación con una posterior sedimentación y desinfección; por último el agua tratada se utiliza con propósitos de riego de áreas verdes de la misma institución.

El propósito de este artículo es llevar a cabo pruebas de tratabilidad (pruebas de jarras) que permitan establecer a través de un análisis estadístico, un modelo que se ajuste a las variables de interés.

La calidad de las aguas ha sido siempre referida a sus características físicas, tales como la turbidez, la presencia de sólidos y características químicas como la demanda química de oxígeno (DQO), la alcalinidad, dureza, conductividad, etc. Bajo condiciones naturales, esto es propiciado por la presencia de coloides inorgánicos y por un alto contenido de materia iónica disuelta en las aguas [1]. Con la finalidad de pulir la calidad del agua para diferentes usos del agua (potabilización, procesos industriales, irrigación, renovación), es necesario desarrollar mecanismos de remoción que permitan el retiro de dicha masa.

Los estudios para la tratabilidad conducen a la utilización de la llamada "Prueba de Jarras", la cual sirve específicamente para probar la efectividad de los coagulantes y de esta manera verificar cual es el mejor tratamiento. La prueba consiste en colocar seis vasos de precipitado conteniendo 500 ml de muestra cada uno, estos vasos se colocan en un equipo que tiene seis paletas de agitación y estas pueden accionarse a diferentes revoluciones por minuto [2].

Para la remoción de las partículas coloidales es necesario desestabilizar el medio mediante la adición de coagulantes químicos (coagulación), y propiciar agitación controlada al agua para lograr la aglutinación de los coloides (floculación), para que finalmente dichas partículas aglutinadas (floculos), se encuentren bajo el efecto de la gravedad (sedimentación) [3]

Metodología

Entre las variables analizadas se encuentran:

SST: esta variable representa los sólidos suspendidos totales que se encuentran en el agua residual al final del tratamiento es decir, aquellos sólidos remanentes que no lograron eliminarse con la aplicación del coagulante. Mientras menor sea esta variable más eficaz es la remoción. [4].

Turbidez: esta variable representa las impurezas contenidas en el agua al final del tratamiento, es una medida de la opacidad. [5].
DQO: la demanda química de oxígeno representa la cantidad de oxígeno necesario para degradar la materia orgánica contenida en el agua residual, mientras menor sea esta variable más eficaz es la remoción. [5].

Conductividad: Esta medición resulta de utilidad para evaluar las variaciones en la concentración de minerales disueltos en el agua residual y el agua tratada.

Dureza: la medición de la dureza por calcio es de importancia técnica debido a los problemas que puede ocasionar al generar incrustaciones en el equipo y tuberías. [5].

Alcalinidad: El conocimiento y el control de este parámetro es importante en el proceso de tratamiento de aguas, sobre todo al tratarse de un proceso de coagulación química pues la alcalinidad posee la propiedad de inhibir la acción del coagulante. [5].

Adicionalmente se midieron parámetros tales como: dureza por calcio, alcalinidad, conductividad, temperatura, el tiempo de agitación y la velocidad de agitación.

Debido a que las aguas negras de la institución están conformadas por agua residual de los servicios sanitarios y de los diversos puntos de descarga de los servicios alimentarios, la carga orgánica contenida se ha ido incrementando con el paso del tiempo. En la Tabla 1 se observan los parámetros medidos y los rangos obtenidos de cada uno.

Parámetros	Rangos
Flujo volumétrico	432.88 a 813.21 m ³ /día
Sólidos totales	755 a 1280 mg/L
SST	5 a 370 mg/L
Turbidez	14.6 a 256 UTN
pH	6.15 a 8.69 unidades de pH
Alcalinidad	166 a 760 mg/L de CaCO ₃
DBO	11.5 a 450 mg/L
DQO	70.18 a 592.35 mg/L
Conductividad	764 a 2170 S/cm
Calcio	101.6 a 225.6 mg/L

Tabla 1 Rangos mínimos y máximos de parámetros en el influente de la planta de tratamiento.

De acuerdo a la caracterización de la Tabla 1, se seleccionaron las siguientes variables en los clarificados:

$Y = DQO = \text{demanda química de oxígeno}$

$X1 = SST = \text{sólidos suspendidos totales}$

$X2 = \text{Turbidez (UTN)}$

$X3 = \text{Dureza por calcio}$

$X4 = \text{Conductividad}$

$X5 = \text{Alcalinidad}$

Se realizaron diez pruebas de jarras de las cuales se extrajeron seis muestras de cada una para dar un total de 60 muestras.

En cuanto al análisis estadístico, se evaluó un modelo de regresión lineal múltiple con aquellas variables que resultaron más significativas en una inspección inicial.

Resultados

Los datos se analizaron con el programa SPSS, primeramente se analizaron las correlaciones entre todas las variables para identificar aquellas que pudieran resultar más significativas para el modelo como se muestra en la Figura 1. La variable turbidez fue la que mostro la mayor correlación con la variable dependiente DQO de 0.753, y las variables SST, dureza y conductividad mostraron correlaciones de 0.353, 0.528 y -0.427 respectivamente. Debido a lo anterior se graficó un diagrama de dispersión matricial con dichas variables. Se puede observar en la Figura 2 que hay una tendencia lineal entre la variable DQO y las variables turbidez y SST que son las que tuvieron los valores de correlación más altos.

		DQO	SST	TURBIDEZ	Alcalinidad	Dureza (Ca)	Conductividad
DQO	Pearson Correlation	1	.353**	.753**	.033	.528**	-.427**
	Sig. (2-tailed)	.	.006	.000	.800	.000	.001
	N	60	60	60	60	60	60
SST	Pearson Correlation	.353**	1	.290*	.188	.273*	-.240
	Sig. (2-tailed)	.006	.	.025	.151	.035	.065
	N	60	60	60	60	60	60
TURBIDEZ	Pearson Correlation	.753**	.290*	1	.102	.615**	-.410**
	Sig. (2-tailed)	.000	.025	.	.436	.000	.001
	N	60	60	60	60	60	60
Alcalinidad	Pearson Correlation	.033	.188	.102	1	-.264*	-.703*
	Sig. (2-tailed)	.800	.151	.436	.	.042	.000
	N	60	60	60	60	60	60
Dureza (Ca)	Pearson Correlation	.528**	.273*	.615**	-.264*	1	-.064
	Sig. (2-tailed)	.000	.035	.000	.042	.	.629
	N	60	60	60	60	60	60
Conductividad	Pearson Correlation	-.427**	-.240	-.410**	-.703*	-.064	1
	Sig. (2-tailed)	.001	.065	.001	.000	.629	.
	N	60	60	60	60	60	60

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Figura 1 Resultados de las correlaciones entre todas las variables

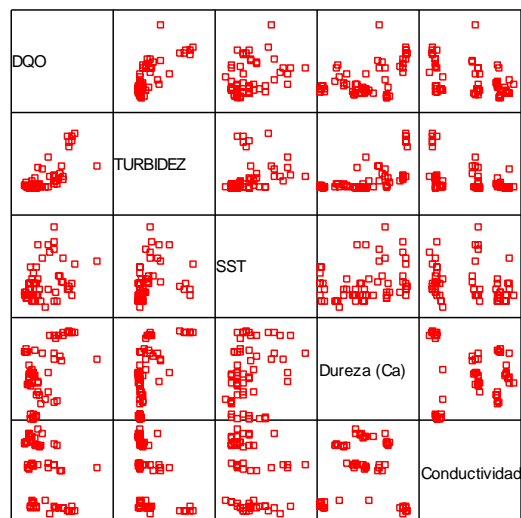


Figura 2 Diagrama de dispersión

En base esta inspección inicial se verificó si la regresión con todas las variables pudiera ser significativa y se obtuvieron los resultados de la Tabla 2. Se planteó la prueba de hipótesis expresada en (1) y (2).

$$H_0: B_1 = B_2 = B_3 = B_4 = 0 \quad (1)$$

$$H_1: B_i \neq 0 \text{ para algún } 1 \leq i \leq 4 \quad (2)$$

Debido a que el p-valor fue de $0.000 < 0.05$, se rechazó la H_0 y por lo tanto la regresión es significativa al juntar todas las variables de interés; sin embargo al analizar los coeficientes de cada una por separado sus p-valoros correspondientes indican que por sí solas no tienen significancia a excepción de la variable turbidez que tuvo un p-valor de 0.000 como se ve en la Tabla 2.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.781 ^a	.610	.582	144,569

a. Predictors: (Constant), Conductividad, Dureza (Ca), SST, TURBIDEZ

b. Dependent Variable: DQO

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1797083	4	449270,781	21,496	,000 ^a
	Residual	1149508	55	20900,143		
	Total	2946591	59			

a. Predictors: (Constant), Conductividad, Dureza (Ca), SST, TURBIDEZ

b. Dependent Variable: DQO

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	135,264	56,271		2,404	,020
	SST	,801	,642	,112	1,247	,218
	TURBIDEZ	2,189	,463	,574	4,733	,000
	Dureza (Ca)	,176	,148	,134	1,193	,238
	Conductividad	-7,41E-02	,046	-,156	-1,604	,114

a. Dependent Variable: DQO

Tabla 2 Resultados del ANOVA para el modelo completo

Como las variables dureza por calcio y conductividad tuvieron valores de correlaciones relativamente bajos se analizó su contribución en la regresión, dado que las otras variables ya estuvieran consideradas en el modelo, y se planteó la prueba de hipótesis expresada en (3) y (4).

$$H_0: B_3 = B_4 = 0 \quad (3)$$

$$H_1: B_3 \neq 0 \text{ ó } B_4 \neq 0 \quad (4)$$

Se obtuvieron los resultados que se muestran en la Tabla 3 a partir del modelo reducido.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,766 ^a	,587	,573	146,034

a. Predictors: (Constant), TURBIDEZ, SST

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1731012	2	865506,107	40,585	,000 ^a
	Residual	1215579	57	21325,943		
	Total	2946591	59			

a. Predictors: (Constant), TURBIDEZ, SST

b. Dependent Variable: DQO

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	94,304	31,831		2,963	,004
	SST	1,047	,633	,147	1,662	,104
	TURBIDEZ	2,709	,339	,711	7,998	,000

a. Dependent Variable: DQO

Tabla 3 Resultados del ANOVA para el modelo reducido

Con los datos del modelo reducido y el completo se verificó la significancia de las variables dureza y conductividad para el modelo.

$$Y = B_0 + B_1SST + B_2 \text{ turbidez} + B_3 \text{ Dureza} + B_4 \text{ Conductividad} \quad (5)$$

$$Y = B_0 + B_1SST + B_2 \text{ turbidez} \quad (6)$$

$$F = \frac{(SSE_r - SSE_c)/(p-m)}{SSE_c/(n-p-1)} = 1.5806 \quad (7)$$

Con el criterio de $F_{0,05,2,55} = 3.15$, 1.5806 no es mayor y por lo tanto no se rechaza la H_0 , por lo que no hay evidencia para decir que la dureza y la conductividad contribuyan significativamente en el modelo.

Posteriormente se planteó si las variables turbidez y dureza contribuyen significativamente, dado que la variable SST está considerada en el modelo y se planteó la prueba de hipótesis como se indica en (8) y (9).

$$H_0: B_2 = B_3 = 0 \quad (8)$$

$$H_1: B_2 \neq 0 \text{ ó } B_3 \neq 0 \quad (9)$$

Se obtuvieron los resultados que se muestran en la Tabla 4 y 5 a partir del modelo completo y el reducido respectivamente.

$$Y = B_0 + B_1SST + B_2 \text{ turbidez} + B_3 \text{ Dureza} \quad (10)$$

$$Y = B_0 + B_1SST \quad (11)$$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,769 ^a	,592	,570	146,584

a. Predictors: (Constant), Dureza (Ca), SST, TURBIDEZ

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1743323	3	581107,680	27,045	,000 ^a
	Residual	1203268	56	21486,928		
	Total	2946591	59			

a. Predictors: (Constant), Dureza (Ca), SST, TURBIDEZ

b. Dependent Variable: DQO

Tabla 4 Resultados del ANOVA para el modelo completo

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,353 ^a	,125	,109	210,895

a. Predictors: (Constant), SST

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	368941,1	1	368941,096	8,250	,006 ^a
	Residual	2579650	58	44476,722		
	Total	2946591	59			

a. Predictors: (Constant), SST

b. Dependent Variable: DQO

Tabla 5 Resultados del ANOVA para el modelo reducido

Con los datos del modelo reducido y el completo se verificó la significancia de las variables turbidez y dureza para el modelo.

$$F = \frac{(SSE_r - SSE_c)/(p-m)}{SSE_c/(n-p-1)} = 32.028 \quad (12)$$

Con el criterio de $F_{0.05,2,56} = 3.15$, el valor de 32.028 es mucho mayor y por lo tanto se rechaza la H_0 , por lo que hay evidencia de que al menos una de las variables (turbidez y/o dureza) contribuye significativamente en el modelo. Por último se realizó una selección hacia atrás; los resultados se muestran en la Tabla 6. Las variables excluidas fueron todas a excepción de la turbidez que tuvo un p-valor de 0.000, con lo cual se demuestra su significancia. Para validar el análisis, se verificó la homocedasticidad así como la normalidad de los residuos la cual se presenta en la Figura 3.

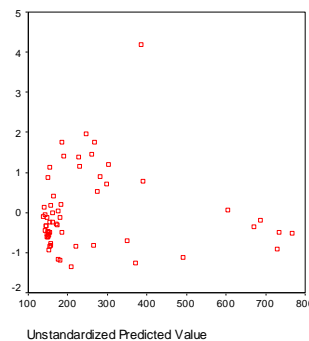


Figura 3. Gráfico de dispersión de los residuos vs valores predichos

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	TURBIDEZ		Forward (Criterion: Probabilty-of-F-to-enter <= ,050)

a. Dependent Variable: DQO

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1672779	1	1672779,019	76,166	,000 ^a
	Residual	1273812	58	21962,275		
	Total	2946591	59			

a. Predictors: (Constant), TURBIDEZ

b. Dependent Variable: DQO

Excluded Variables^b

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	SST	,147 ^a	1,662	,104	,214	,916
	Dureza (Ca)	,104 ^a	,947	,348	,124	,622
	Conductividad	-,142 ^a	-1,516	,135	-,197	,832
	Alcalinidad	-,044 ^a	-,507	,614	-,067	,989

a. Predictors in the Model: (Constant), TURBIDEZ

b. Dependent Variable: DQO

Tabla 6 Regresión lineal con selección hacia atrás

En la Figura 3 se observa que los residuos estandarizados se distribuyen a los largo del grafico de dispersión sin seguir ningún patrón específico, validando así la homocedasticidad.

En la tabla no.7 se observa el p-valor para los residuos del modelo con una variable (turbidez) y como $0.088 > 0.05$ no se rechaza H_0 y no hay evidencia en contra de la normalidad de los residuos. Adicionalmente se elaboró un gráfico Q-Q para verificar normalidad, el cual presenta en la Figura 4.

One-Sample Kolmogorov-Smirnov Test

		Standardized Residual
N		60
Normal Parameters ^{a,b}	Mean	,0000000
	Std. Deviation	,99148923
Most Extreme Differences	Absolute	,161
	Positive	,161
	Negative	-,088
Kolmogorov-Smirnov Z		1,249
Asymp. Sig. (2-tailed)		,088

a. Test distribution is Normal.

b. Calculated from data.

Tabla 7 Prueba de kolmogorov-smirnov para la normalidad

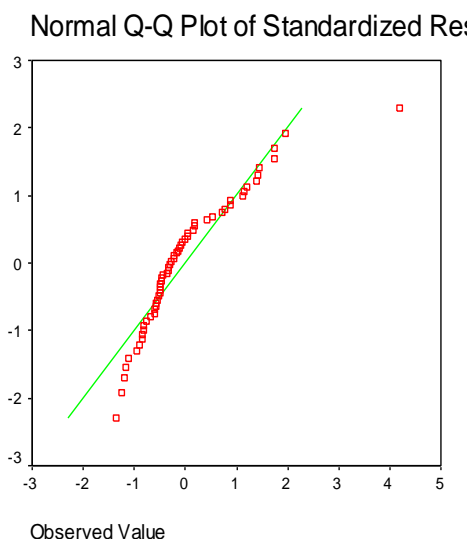


Figura 4 Gráfica en papel de probabilidad normal

Conclusiones

El modelo de regresión lineal que mejor describe los datos analizados en los clarificados provenientes de las diez pruebas de jarras, es un modelo simple donde se incluye la variable turbidez:

$$Y (\text{DQO}) = B_0 + B_1 \text{Turbidez}$$

El resultado coincide con lo descrito en la teoría de tratabilidad de aguas residuales, pues se tiene evidencia de que la turbidez es un factor preponderante que influye sobremanera en la demanda química de oxígeno. Se sabe que los sólidos suspendidos totales están relacionados con las lecturas de turbidez obtenidas en laboratorio, en este estudio se pudo verificar que los SST fueron la segunda variable de importancia aparente al tener una correlación de 0.353, aunque por ser una correlación débil esto no se reflejó significativamente en el modelo final.

Referencias

AWWA Coagulation Committee. "Committee Report: Coagulation as an Integrated Water Treatment Process". Journal of AWWA 81, pp. 72-78. 1989.

APHA-AWWA-WPCF, "Métodos normalizados para el análisis de aguas potables y residuales". Ed. Díaz de Santos, 17ª edición. 1989.

Dentel, S.K., Gossett, J.M. "Mechanisms of Coagulation with Aluminum Salts". Journal of AWWA. 80, pp. 187-198. 1988.

Crites, R., Tchobanoglous, G. "Tratamiento de aguas residuales en pequeñas poblaciones". Ed. Mc. Graw-Hill. México, 2000.

Eckenfelder, W. "Industrial water pollution control". Ed. Mc. Graw-Hill, 3ra. Ed. 2000.