



Universidad del Azuay

Facultad de Ciencia y Tecnología

Escuela de Biología del Medio Ambiente

**Determinación de la Demanda Química de Oxígeno en
efluentes de Industrias Lácteas empleando un método
amigable con el Medio Ambiente**

**Trabajo Previo a la Obtención del Título de
Biólogo**

Autor:

Luis Fernando Cárdenas Calle

Director:

Dr. Piercosimo Tripaldi

**Cuenca, Ecuador
2008**

Dedicatoria

A Dios por su guía y fortaleza
A mis padres por su amor y apoyo constante
A mis hermanos por su comprensión
A mis abuelos por su gran cariño

Agradecimientos

Al Dr. Piero Tripaldi por su guía en la realización de esta tesis. Por su total y generosa ayuda recibida en el campo profesional. Y por la confianza otorgada en el trabajo que desempeño.

A la Ph.D Dr. María Elena Cazar por la ayuda brindada en la realización de mis principales trabajos científicos y por su sincera amistad.

Al Blgo. Edwin Zárate por la confianza que me tuvo al confiarme la realización de importantes trabajos.

A los Biólogos Sebastián Padrón y Gonzalo Sotomayor por todos los valiosos consejos dados.

Al Ing. Cristian Rojas por todo el apoyo que me ha brindado en mi ejercicio laboral.

RESUMEN

Determinación de la Demanda Química de Oxígeno en efluentes de Industrias Lácteas empleando un método amigable con el Medio Ambiente

El parámetro más utilizado para determinar el grado de afectación al agua por las Industrias Lácteas es la determinación de la DQO. El método tradicional presenta varias deficiencias, por lo que es de gran importancia la búsqueda de nuevas metodologías para la eficiente determinación de este parámetro. Se evaluó la DQO clásico y el espectro de bandas ultravioletas sobre las aguas residuales de la Planta de Lácteos. Se construyó un modelo mediante algoritmos genéticos el cual determinó que existe una correlación entre el DQO clásico y el espectro de bandas ultravioletas. Con este método se puede minimizar las deficiencias que presenta el método clásico de análisis.

ABSTRACT**Determination of the Chemical Oxygen Demand in effluent from Dairy Industries using a method friendly Environment**

The most commonly parameter used to determine the degree of impact of water by the dairy industries is the COD determination. The traditional method has several shortcomings. For this reason is very important to search for new methodologies oriented to the efficient determination of this parameter. We evaluated the COD by the classic method and the ultraviolet spectrum bands on the wastewater from a dairy plant. A model using genetic algorithms was developed, which determined a correlation between the COD determined by the classic method and the ultraviolet spectrum bands. This approach can minimize the weakness presented in the classic method of analysis.

OBJETIVOS

OBJETIVO GENERAL

Mediante el presente estudio se tiene como objetivo determinar si existe una correlación válida entre las señales espectro-fotométricas y la determinación de la Demanda Química de Oxígeno de los desechos de las Industrias Lácteas. Lo cual sería de gran importancia ya que el método tendría varias ventajas con relación al método clásico.

OBJETIVOS ESPECIFICOS

- Obtener datos comparables con el método clásico.
- Realizar una Validación Interna del método.
- Establecer el costo del método utilizado y compararlo con el costo de la técnica clásica.
- Establecer un monitoreo ambiental a bajo costo de la Planta de Lácteos.

Índice de Contenidos

Dedicatoria.....	ii
Agradecimientos.....	iii
Resumen.....	iv
Abstract.....	v
Objetivos.....	vi
Índice de Contenidos.....	vii
Índice de Ilustraciones y Cuadros.....	viii
Tabla de Abreviaturas.....	xi
Introducción.....	1
CAPITULO I: METODOLOGIA	
1.1 Zona de Estudio.....	12
1.2 Toma de Muestras.....	13
1.3 Trabajo de Laboratorio.....	14
1.4 Recolección y Análisis de lo Datos.....	15
CAPITULO II: RESULTADOS	
2.1 Modelo Matemático obtenido.....	16
2.2 Parámetros de Calidad del Modelo.....	17
2.3 Validación Interna del Método.....	18
2.4 Validación Externa del Método.....	20
2.5 Monitoreo de la Planta de Lácteos.....	21
2.6 Costo del Método.....	22
2.7 Beneficios al Medio Ambiente.....	23
CAPITULO III: DISCUSION.....	25
CAPITULO IV CONCLUSIONES.....	29
Referencias Bibliográficas.....	31
Anexo 1: Fotografía de la Planta de Lácteos.....	33
Anexo 2: Fotografía de la Producción de Quesos en la Planta.....	34
Anexo 3: Fotografía Espectrofotómetro UV-vis.....	35
Anexo 4: Fotografía del registrador de datos.....	36
Anexo 5: Fotografía Muestras y digestor Kjeldahl.....	37
Anexo 6: Matriz de Datos n x p utilizada para el cálculo del modelo matemático.....	38

Índice de Ilustraciones y Cuadros

Tabla 1: Parámetros de calidad del modelo matemático.....	17
Tabla 2: Parámetros encontrados en la validación interna del método.....	20
Figura 1: Representación de la optimización de la respuesta mediante los Algoritmos Genéticos.....	8
Figura 2: Planta Piloto de Lácteos de la Escuela de Ingeniería de Alimentos.....	12
Figura 3: Canales donde se recolectaron las muestras para el análisis de la DQO.....	13
Figura 4: Muestras de DQO clásico.....	14
Figura 5: Ejemplo de algunos Espectros de las muestras.....	16
Figura 6: Relación entre DQO experimental y los valores calculados en interpolación y predicción por el modelo propuesto.....	17
Figura 7: Gráfico de Williams del modelo propuesto.....	18
Figura 8: Marmita de capacidad de 100L para la elaboración de quesos.....	22

Tabla de Abreviaturas

Abreviatura	Correspondencia
DQO	Demanda Química de Oxígeno
DBO	Demanda Bioquímica de Oxígeno
nm	Nanómetros
N	Normalidad
L	Litros

**Cárdenas Calle Luis Fernando, Trabajo de Graduación
Dr. Piercosimo Tripaldi
Abril 2008**

**DETERMINACIÓN DE LA DEMANDA QUÍMICA DE OXÍGENO EN
EFLUENTES DE INDUSTRIAS DE LÁCTEOS EMPLEANDO UN MÉTODO
AMIGABLE CON EL MEDIO AMBIENTE**

INTRODUCCION

En los últimos años el recrudecimiento de los problemas ambientales relacionados con las actividades humanas han creado una gran presión sobre los recursos naturales en especial sobre el agua. El agua a pesar de constituir el 75% de la superficie de nuestro planeta solamente un 2.5% es dulce y de ésta el 0.94% está disponible para el consumo humano. La mala distribución y consumo excesivo del agua han producido graves problemas de contaminación. En la industria se utiliza gran cantidad de agua que es contaminada dependiendo de la industria y no es recuperada (Glynn, 1999).

La industria láctea consume gran cantidad de agua la misma que es utilizada en todos los procesos de fabricación de los diferentes productos. En la mayoría de las industrias lácteas del país se fabrican 6 productos: obtención de leche fresca, pasteurizada y descremada, la fabricación de quesos, fabricación de mantequillas, fabricación de leche condensada, la fabricación de leche en polvo y la fabricación de yogurt. En la planta piloto de lácteos de la Universidad del Azuay se realizan dos procesos: fabricación de quesos y yogurt (Álvarez, 2006).

En el Ecuador el requerimiento de productos lácteos al año por persona es de 110 litros. Cada día en el país se producen 4,4 millones de litros de los cuales el 42% son empleados por la industria, el resto corresponde a la leche no procesada que se vende sobre todo en el área rural del país. La producción industrial destina el 41% de su producción a la leche en funda, el 24% a quesos, el 15% a leche en cartón, el 11% a leche en polvo, el 7% a yogurt y el 2% a otros (Álvarez, 2006).

Estos altos requerimientos diarios de la sociedad ecuatoriana, han incrementado en gran manera los problemas ambientales relacionados con la industria láctea sobre todo con el recurso agua. Las aguas residuales de las industrias de lácteos se originan principalmente en el proceso de lavado y limpieza de las tuberías, en los recipientes que transportan la leche y en otros equipos; los derrames de leche por fugas, sobre-flujos, mal funcionamiento de equipos. Además, el proceso de desechar los subproductos de la elaboración de quesos y mantequillas tales como el suero de la leche. El 94% de los desechos líquidos se originan en los procesos de producción ya sea de leche de consumo diario o en la producción de quesos y mantequilla, el 6% se origina en los desechos de limpieza de equipos e instalaciones (Fundación Natura, 1999).

Para la evaluación del potencial contaminante de estos desechos, tres son los más importantes: Demanda Química de Oxígeno (DQO), la demanda Bioquímica de Oxígeno (DBO) y los sólidos en suspensión totales. Si los desechos líquidos del procesamiento industrial de la leche se descargan sin ningún tratamiento a un río o lago, estos desechos sufren una biodegradación con el consiguiente consumo de oxígeno del agua. Esto puede causar la muerte de organismos sensibles, también se pueden producir condiciones anaeróbicas con la presencia de malos olores, y eutrofización del medio receptor (Fundación Natura, 1999).

La calidad de un paisaje puede ser perjudicada por el desplazamiento de un complejo industrial y la aparición de sustancias contaminantes, esto podría perjudicar la calidad estética de un río o lago. Los materiales en suspensión disminuyen el paso de luz a través del agua esto retarda el crecimiento de flora acuática de la cual pueden depender animales para su alimentación. Otro efecto de los sólidos en suspensión es que se produce sedimentación que arrastra los nutrientes al fondo del cauce donde se depositan (Fundación Natura, 1999).

La composición de los residuos de las industrias lácteas tiene las siguientes características: contiene materia orgánica y otros productos susceptibles a descomposición que originan ácido láctico, precipitan la caseína y otros compuestos nitrogenados. El pH es bajo entre 4.5 y 4.8, la composición media del vertido es el siguiente: materia orgánica 63%, nitrógeno 75%, fósforo 7%, potasio 0.7%. Las aguas residuales de este tipo de industria no varían en gran mayoría su composición a

excepción de la fabricación de quesos, son aguas ácidas por la presencia de suero y son las que más sólidos en suspensión presentan debido al cuajo. El resto de productos presenta vertidos neutros o ligeramente alcalinos con tendencia a ser ácidos por la fermentación del azúcar para transformarse en ácido láctico (Sans, 1999).

Es de gran importancia las técnicas cuantitativas de análisis para la caracterización de estos residuos, las mediciones cuantitativas de cualquier clase son la piedra angular para su caracterización. Cada residuo a ser analizado debe tener metodologías definidas para su determinación, este análisis necesita el uso de métodos, procedimientos analíticos de laboratorio y de campo, que hayan probado su confiabilidad al ser usados por muchas personas y con gran variedad de materiales. Por lo dicho anteriormente el parámetro más utilizado y que mejores resultados presenta para determinar el grado de afectación al agua por este tipo de industria es la determinación de la DQO en las aguas de desecho. La prueba de la DQO es ampliamente usada como una forma de medir la concentración de la materia orgánica en los residuos domésticos e industriales. (Sawyer, 2001)

Esta prueba permite medir en un residuo la cantidad total de oxígeno que se requiere para la oxidación de la materia orgánica a dióxido de carbono y agua. La prueba se basa en que todos los compuestos orgánicos, con pocas excepciones, pueden ser oxidados por la acción de agentes oxidantes fuertes en condiciones ácidas. Durante la determinación de la DQO, la materia orgánica es convertida a dióxido de carbono y agua independientemente de la capacidad biológica de las sustancias para ser asimiladas. Por ejemplo: la glucosa y la lignina son completamente oxidadas, en consecuencia, los valores de DQO son mayores que la DBO y pueden ser aún mayores cuando existen cantidades significativas de materia orgánica biológicamente resistente. Una de las principales limitaciones de la prueba de la DQO es la imposibilidad para diferenciar entre la materia biológicamente oxidable y la materia orgánica biológicamente inerte (Sawyer, 2001).

Una de las ventajas que proporciona la prueba de la DQO es el poco tiempo de evaluación del análisis en relación al DBO que necesita de 5 días para su análisis, mientras que la DQO se puede realizar en unas tres horas. Además se pueden inferir los datos de la DQO en DBO después de haber acumulado suficiente experiencia en su

determinación y establecer factores de correlación confiables. Desde hace mucho tiempo los agentes químicos oxidantes han sido usados para medir la demanda de oxígeno de las aguas contaminadas. Durante muchos años se utilizaron soluciones de permanganato de potasio y los resultados se llamaban consumo de oxígeno por permanganato (Sawyer, 2001).

La oxidación producida por el permanganato era altamente variable en los diferentes tipos de compuestos y el grado de oxidación cambiaba de manera considerable con la concentración del reactivo usado. Los valores del consumo de oxígeno eran significativamente menores que los de DBO en cinco días. Este hecho demostraba la incapacidad del permanganato para llevar la oxidación hasta un punto final específico. El sulfato cérico, el yodato de potasio y dicromato de potasio son otros agentes oxidantes que han sido extensamente estudiados para la determinación de la DQO. Se ha visto que el dicromato de potasio es el más práctico de todos puesto que es capaz de oxidar casi completamente una gran variedad de sustancias orgánicas hasta dióxido de carbono y agua (Sawyer, 2001).

Debido a que todos los agentes oxidantes deben ser usados en exceso es necesario medir el exceso que queda al final del periodo de la reacción con el fin de calcular la cantidad realmente utilizada para la oxidación de la materia orgánica. Un importante punto a favor del dicromato de potasio es que el exceso se puede medir con relativa facilidad. Para que el dicromato de potasio oxide completamente la materia orgánica, la solución debe ser fuertemente ácida y estar a temperatura elevada. En consecuencia, los materiales volátiles originalmente presentes y los que se han formado durante el periodo de digestión se pierden a menos que se hagan las provisiones necesarias para evitar el escape de estas sustancias (Sawyer, 2001).

Con este fin se suelen emplear los condensadores de reflujo que permiten que la mezcla haga ebullición sin pérdida significativa de los compuestos orgánicos volátiles. Ciertos compuestos orgánicos, especialmente los ácidos grasos de bajo peso molecular no son oxidados por el dicromato de potasio, a menos que haya un catalizador. Se ha visto que el ión plata es efectivo en este sentido. Los hidrocarburos aromáticos y la piridina no se oxidan bajo ninguna circunstancia. El procedimiento común en la determinación de la DQO se realiza mediante una solución de dicromato de potasio 0.25N y la adición de

ácido sulfúrico concentrado, la cual oxida la materia orgánica presente en la muestra de agua, la misma que es llevada a reflujo durante dos horas (Sawyer, 2001).

Luego se realiza una titulación con sulfato de amonio ferroso para determinar la cantidad de dicromato de potasio consumido y calcular la materia orgánica oxidable en términos de oxígeno equivalente. Esta técnica utiliza gran cantidad de las sustancias mencionadas, por lo cual se realiza un proceso alternativo de análisis en el cual se lleva a reflujo abierto las muestras, y las cantidades de reactivos son mínimas pero teniendo en cuenta una relación 1:1 en la solución de dicromato de potasio y la muestra analizada. El método clásico de análisis presenta varias deficiencias. El procedimiento patrón que utiliza dicromato de potasio en medio ácido a temperaturas altas genera un gran volumen de residuos líquidos peligrosos como ácido, cromo, hierro y plata que son desechados por los desagües, estos residuos pueden llegar a las plantas de tratamiento y a las aguas superficiales (Sawyer, 2001).

Se han desarrollado técnicas alternativas donde se minimizan las cantidades de reactivos utilizados en el análisis, pero no se ve al problema real del análisis que es la utilización de productos nocivos para el medio ambiente, por esta razón es de importancia la búsqueda de técnicas alternativas para determinar la DQO en muestras de importancia ambiental. Al realizar una digestión en medio ácido y a temperaturas elevadas se pueden producir accidentes de laboratorio, la digestión demora dos horas lo cual se transforma en un alto costo energético, la utilización de catalizadores como sulfato de plata hacen que el análisis tenga un costo alto, acarrea un tiempo de análisis elevada, la reproductividad de la técnica es bastante baja elevando los errores a un 30% (Peralta, 2005).

Esta técnica es empleada en la mayoría de legislaciones ambientales como referencia para evaluar el potencial contaminante de un efluente, por lo cual es de gran importancia este método. La búsqueda de nuevas técnicas que minimicen la contaminación al no utilizar productos químicos en el análisis de la muestra, reducir el error de la prueba, minimizar accidentes en el laboratorio, reducir el gasto energético y sobre todo disminuir el costo del análisis. Esto ha incentivado la búsqueda de nuevas técnicas que sean amigables con el medio ambiente y que tengan los mismos resultados que la técnica clásica de determinación de la DQO (Peralta, 2005).

Se ha establecido que existe una alta correlación entre las señales espectro-fotométricas y la DQO para diversos residuos industriales. Principalmente considerando el área espectral de los 200 a los 400 nm. Con estos datos del área espectral se puede determinar las bandas más significativas las cuales servirán para la determinación de la DQO. Las ventajas del método es que se reduce la utilización de compuestos químicos perjudiciales para el medio ambiente, se reduce el error de la prueba a un 10% debido a la poca manipulación de la muestra, el gasto energético es casi nulo, el costo del análisis es bajo (Peralta, 2005).

Se puede evaluar la eficiencia de este método gracias a las técnicas de análisis multivariante de datos, estas técnicas son un conjunto de métodos estadísticos cuya finalidad es analizar simultáneamente conjuntos de datos en donde existen una serie de variables medidas para cada individuo u objetos estudiados. Por esta razón, este tipo de análisis da un mejor entendimiento de la problemática del fenómeno estudiado obteniendo toda la información posible de los datos estudiados, ya que con técnicas univariantes o bivariantes no se lograría este propósito. Además el análisis ayuda al investigador a tomar decisiones óptimas para solucionar un problema propuesto teniendo en cuenta toda la información disponible del conjunto de datos analizados (Frank, 1994).

Para llevar a cabo un análisis de este tipo se deben considerar los siguientes pasos:

- 1) Establecer los objetivos del análisis.
- 2) Diseñar el análisis.
- 3) Evaluar las hipótesis.
- 4) Realizar el análisis.
- 5) Interpretar los resultados obtenidos.
- 6) Validar los resultados.

El análisis consiste en la búsqueda de una correlación válida entre las diferentes absorbancias y la DQO de muestras de las industrias lácteas por lo que el análisis se basa en una regresión multivariante. La regresión lineal utiliza una ecuación matemática para modelar la relación entre las cantidades medidas y observadas en los datos proporcionados por el análisis. El objetivo de este análisis es doble: la modelación y la

predicción de nuevos objetos introducidos en el modelo. La relación se describe en forma algebraica como:

$$y = f(x) + e$$

Donde x son las variables predictoras y y son las respuestas encontradas para dicha variable, $f(x)$ es la parte sistemática, y e es el error del modelo matemático encontrado. Para obtener una regresión de buena calidad se ha realizado primero la selección de las mejores variables mediante algoritmos genéticos. Este es un nuevo campo en rápido desarrollo y que tiene un gran potencial en aplicaciones de análisis de datos. Los algoritmos genéticos son una potente estrategia de investigación quimiométrica aplicada en problemas de optimización de larga escala. Este método fue propuesto por Holland en 1975, y en estos últimos tiempos han sido muy utilizados para la solución de problemas de optimización en muchos campos como: el análisis de las imágenes, reconocimiento de modelos matemáticos, métodos de modelación, robótica, proyectos de redes e investigación operativa (Delgado, 1998).

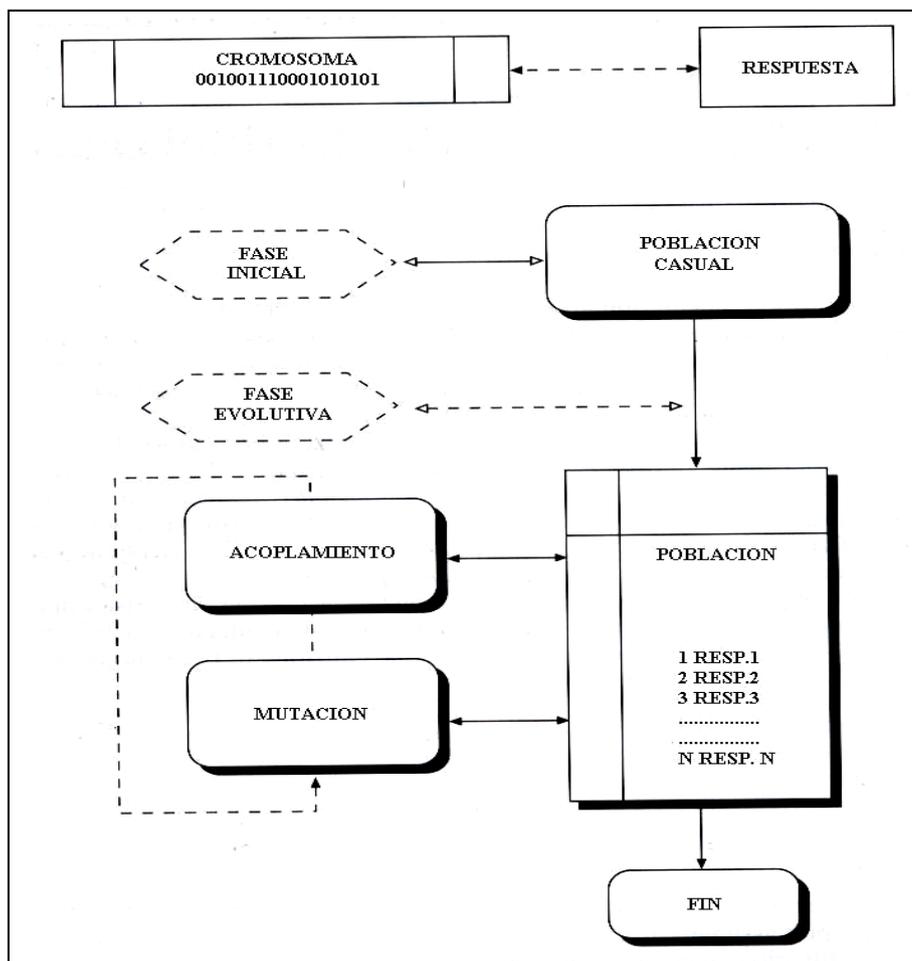
Los Algoritmos Genéticos son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización de sistemas. Se basan en el proceso genético evolutivo de los organismos vivos. Un algoritmo genético consiste en un ciclo interactivo de creaciones de conjuntos (cromosomas) de variables originales que generan poblaciones las cuales son las óptimas para la construcción del modelo predictor. Estos conjuntos producen modelos de regresión OLS de los cuales se evalúa la calidad generalmente mediante el R^2_{cv} . Son retenidos los mejores y sobre esta población se procede al apareamiento de cromosomas y a mutaciones para producir nuevos conjuntos predictores. Se retienen siempre los mejores y se repite el ciclo interactivamente hasta alcanzar la calidad de la regresión deseada. En los algoritmos genéticos se simula el proceso de selección natural y supervivencia en donde los más aptos sobreviven. De esta manera se obtiene un conjunto óptimo de variables base de la regresión (Massart et al, 1997).

La base para la aplicación de los algoritmos genéticos es la descripción de la presencia de cada variable en el modelo mediante un código binario, donde cada término binario constituye un bit, que tiene un número igual a 0 ó 1. En el lenguaje de los algoritmos

genéticos, cada codificación binaria de un número (bit) constituye un gen, la concatenación de grupos de bit (genes) constituye un cromosoma. Cada cromosoma es una representación de un punto en el espacio p-dimensional de los parámetros independientes de optimización, donde cada parámetro es representado por un gen (Leardi, 2001).

La estrategia para la optimización de la función respuesta es representada en la siguiente figura:

Figura 1: Representación de la optimización de la respuesta mediante Algoritmos Genéticos (modificado de Todeschini, 1998. pag: 172)



Fuente: Todeschini, 1998. pag: 172

La estrategia comienza con una población artificial de individuos que para el presente estudio son la selección de las variables utilizadas. Una vez que está prefijada la

dimensión de la población de cromosomas comienza la fase inicial donde la población viene generada utilizando la mejor N respuesta obtenida con valores de los parámetros utilizados. Sucesivamente se pasa a la fase evolutiva en donde se da un proceso doble, en donde se combinan (acoplan) los cromosomas presentes en la población, alternativamente e interactivamente se producen mutaciones casuales. En ambos procesos para cada nuevo caso obtenido viene calculada la respuesta y si ésta es mejor que la que está presente actualmente en la población, el cromosoma entra a formar parte de la población sustituyendo al peor. La fase evolutiva pone su fin si se satisface el criterio que se está optimizando, en este caso el Q^2 , finalizando el algoritmo (Todeschini, 1998).

El diverso uso de los algoritmos genéticos tiene su mayor interés en la aplicación de búsqueda de las mejores variables que optimizan un modelo de clasificación o de regresión. Un gen está constituido de un 1 ó 0, que son valores que representan la presencia o ausencia de la variable en el modelo. Se presenta el siguiente ejemplo de 10 variables:

Variable	1	2	3	4	5	6	7	8	9	10
Valor binario	0	1	1	0	1	0	0	1	1	0

En el gráfico se indica la presencia del modelo en la variable 2, 3, 5, 8 y 9 mientras que las variables restantes son momentáneamente excluidas. Estos conjuntos producen modelos de regresión OLS de los cuales se evalúa la calidad generalmente mediante el R^2_{cv} . Son retenidos los mejores y sobre esta población se procede al apareamiento de cromosomas y mutaciones para producir nuevos conjuntos predictores. Se retienen siempre los mejores y se repite el ciclo interactivamente hasta alcanzar la calidad de la regresión deseada. En el caso de un modelo de regresión, el parámetro a maximizar es el Q^2 (el porcentaje de varianza explicada en predicción) mientras en un caso de clasificación, se minimiza el $MR_{cv\%}$ (“misclassification risk” en predicción) (Todeschini, 1998).

El método multivariante de mínimos cuadrados es un modelo de regresión lineal que evalúa los parámetros y las respuestas, los coeficientes b de la regresión se calculan resolviendo la ecuación de los mínimos cuadrados que es la siguiente:

$$y = f(x) + e$$

Donde x son las variables predictoras, (y) son las respuestas encontradas para dicha variable, y e es el error del modelo matemático encontrado.

$$X^T y = X^T X b$$

Luego se multiplica por X^T tanto la parte derecha como la izquierda de la ecuación.

$$(X^T X)^{-1} X^T y = (X^T X)^{-1} X^T X b$$

Donde $(X^T X)^{-1} X^T X = I$, la solución resulta:

$$b_{ols} = (X^T X)^{-1} X^T y$$

Donde:

b = vector de los coeficientes del modelo

X = matriz de las variables predictoras (absorbancia normalizada a las longitudes de onda escogidas por el algoritmo genético)

y = respuestas experimentales (Valores de la DQO obtenidos)

(Draper, 1966)

Donde la expresión anterior se conoce como Mínimos Cuadrados Multivariante.

La varianza de los vectores de los coeficientes es una medida de la estabilidad de los coeficientes del modelo de regresión y se efectúa mediante la siguiente expresión:

$$V(b_{ols}) = \sigma^2 \cdot \text{tr}(X^T X)^{-1} = \sigma^2 \cdot \sum_{j=1}^p (1/\lambda_j)$$

Donde σ^2 es la varianza debida al error experimental, X es la matriz del modelo, λ son los autovalores de la matriz $(X^T X)^{-1}$, (matriz de dispersión del modelo). Se observa que la varianza de los coeficientes (su error) depende del producto de dos contribuciones separadas, el primero (σ^2) que depende solo de la respuesta y el segundo que depende solo del espacio de los predictores. En particular los elementos diagonales principales de $(X^T X)^{-1}$ son proporcionales a la incertidumbre con la cual vienen determinados los parámetros del modelo. La representación de la varianza de los coeficientes en el

espacio ortogonal de la cual λ son los autovalores (último término de la expresión) se puede observar que la eventual presencia de colinealidad entre las variables independientes depende del último auto-valor, si es muy pequeña o casi nulo en este caso la varianza de los coeficientes tienden al infinito, con lo cual se obtiene un modelo que es completamente inestable (Todeschini, 1998).

Estos coeficientes son estimados reduciendo al mínimo la correlación entre la combinación lineal de los predictores y la respuesta. Además mediante esta ecuación se puede establecer la importancia de las variables predictivas en donde se puede determinar los “leverage”, estos están definidos por la matriz H que tiene el nombre de matriz de influencia, que viene definida como:

$$H = (X^T X)^{-1} X^T$$

El valor mínimo de “leverage” se encuentra en correspondencia al centroide del espacio del modelo; mientras el valor máximo es relativo a los puntos más distantes del centro del modelo. Mediante estos parámetros se puede evaluar la calidad del modelo matemático obtenido en la regresión lineal. (Draper, 1966).

Para evaluar la calidad del modelo se determinó R^2 , y además se valoró el poder predictivo del modelo mediante Q^2 , mediante los métodos “*leave-one-out*” (LOO) y “*leave-more-out*” (LMO). El método LOO es un método de validación que perturba en mínima forma al modelo calculado utilizando $n-1$, es decir un objeto viene excluido del modelo, se calcula nuevamente el modelo y las respuestas se van acumulando hasta estimar un parámetro predictivo del modelo. El método LMO es una variante del método LOO donde se sacan un grupo de objetos para que no entren en el cálculo, aquí se debe establecer el porcentaje de objetos que no entrarán en el modelo, luego de esto se calcula el modelo sin el porcentaje de objetos, las respuestas se acumulan y se estima el valor predictivo del modelo (Meloun et al, 1994)

Estos modelos en predicción sirven para encontrar valores de DQO en muestras de industrias lácteas cuyo DQO se desconozca, estos valores serán determinados mediante el modelo matemático encontrado y con lo cual no se contribuirá a la contaminación del agua en el momento del análisis de la muestra.

CAPITULO 1

METODOLOGIA

1.1 Zona de Estudio

El presente estudio se realizó en la Universidad del Azuay, en la Facultad de Ciencia y Tecnología en la Planta Piloto de Lácteos perteneciente a la Escuela de Ingeniería en Alimentos. En la Planta Piloto se procesan queso, manjar de leche y yogurt. La capacidad de la planta es la siguiente: se tiene una yogurtera de 40L de capacidad, una marmita para la producción de quesos de una capacidad de 100L donde se obtienen 13 kilogramos de queso. La planta no posee una producción continua debido a que es una planta dedicada a la Didáctica donde se realizan de dos a tres veces a la semana prácticas docentes.

Figura 2: Planta Piloto de Lácteos de la Escuela de Ingeniería de Alimentos



Elaborado: Fernando Cárdenas Calle

1.2 Toma de Muestras

El objetivo de la toma de muestras es la obtención de una proporción de material cuyo volumen sea lo suficientemente pequeño que pueda ser transportado y manipulado en el laboratorio con facilidad sin que por ello deje de representar con exactitud su contenido. Por lo cual se siguió el manual de recolección de muestras de la APHA donde se describen los métodos estandarizados para la recolección de las muestras. (Sawyer, 2001).

Las muestras fueron tomadas en envases estériles de dos litros de capacidad, las mismas que fueron recolectadas de las canaletas de evacuación de desechos de la planta procesadora de lácteos. Los muestreos se realizaron periódicamente, es decir que estos incluyeron todas las actividades de elaboración de los diferentes productos de la planta de lácteos con lo cual se llegó a tener un monitoreo continuo. Para este fin se obtuvieron 115 muestras las cuales fueron llevadas al laboratorio LAUDA para su posterior análisis.

Figura 3: Canales donde se recolectaron las muestras para el análisis de la DQO



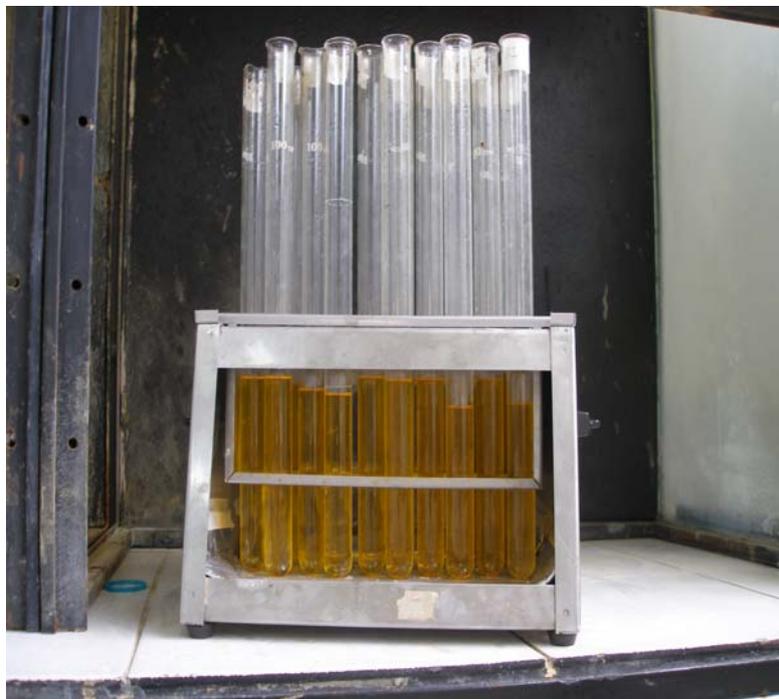
Elaborado: Fernando Cárdenas Calle

1.3 Trabajo de Laboratorio

Las muestras colectadas en la Planta Piloto de Lácteos fueron analizadas en el LAUDA, llevando a cabo el siguiente protocolo:

1. Recolección de la muestra (residuo líquido)
2. Dilución 1:1 con agua destilada
3. Verificación de la dilución en espectrofotómetro. El barrido en UV para esta solución no debe superar las 2 unidades de absorbancia. Si esto ocurre es necesario diluir la muestra.
4. Medición y análisis del espectro de bandas ultravioletas a las siguientes longitudes de onda: 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310nm (200-400nm)
5. Determinación de la DQO mediante el método clásico (APHA, 1998).
6. Digitalización de datos

Figura 4: Muestras DQO clásico



Elaborado: Fernando Cárdenas Calle

1.4 Recolección y Análisis de Datos

Con los datos obtenidos se construye una matriz de datos $n \times p$ donde n son el número de objetos, que para el presente estudio fueron 115 y p el número de variables que son las 12 longitudes de onda y sus diferentes absorbancias. Definida la matriz de datos se auto-escalaron los valores para que su importancia en el modelo de regresión no dependa del tamaño de la longitud de onda para cada muestra. (Ildiko, 1994).

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Con los datos escalados se procedió al control del “leverage” en el programa MATLAB 7.0, los “leverage” son puntos predictivos “X” demasiado altos los cuales pueden influir en la construcción del modelo matemático definitivo. Luego del control de las variables predictivas se realizó la selección de las bandas mas óptimas para la construcción del modelo de regresión entre la absorbancia de las diferentes longitudes de onda de las muestras y su respectivo DQO; este proceso se realizó mediante la utilización de un Algoritmo Genético implementado en el programa MOBY-DIGS que ha sido puesto a disposición por su autor (Todeschini y colaboradores, 2004).

El mejor modelo obtenido ha sido validado mediante “cross-validation” con el 10% de los casos seleccionados al azar y ubicados en el “test set”. Mediante gráfico de Williams en el cual se grafican los residuos estandarizados en predicción “Jackknifed”:

$$r'_{i/i} = \frac{r_i}{s_i \sqrt{1 - h_{ii}}}$$

Se ha controlado la aptitud de las variables predictoras “X” para producir un buen modelo en predicción que permita el ingreso de un nuevo objeto en el cual se pueda calcular el valor de su respectivo DQO mediante la siguiente expresión:

$$h_{i,i} = x_i^T (X^T X)^{-1} x_i$$

El diagnóstico del modelo se ha realizado con el paquete SCAN (Todeschini, 1995).

RESULTADOS

2.1 Modelo Matemático Obtenido

El Algoritmo Genético implementado en MOBY-DIGS seleccionó como mejor conjunto de variables independientes para la regresión a las siguientes longitudes de onda:

220, 260, 270,310nm

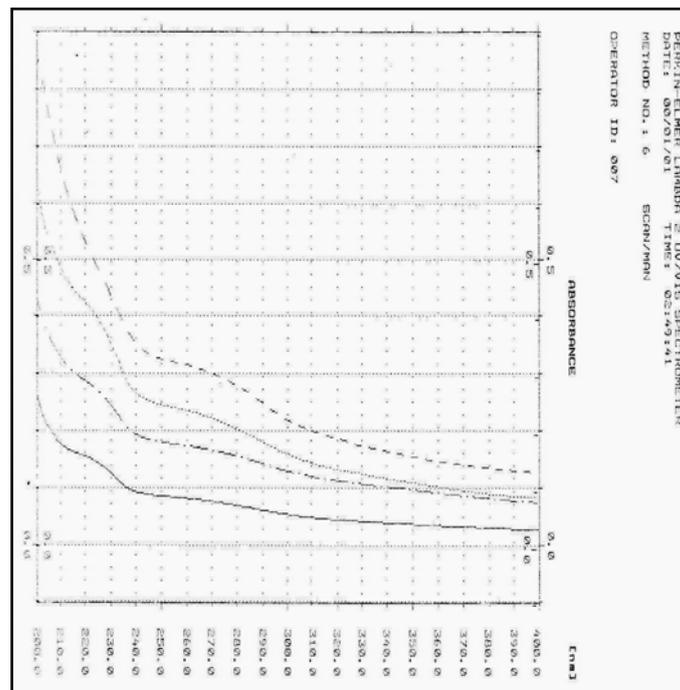
Con las cuales se construyó el siguiente modelo matemático:

$$D.Q.O. = b_0 + b_1\lambda_{220nm} + b_2\lambda_{260nm} + b_3\lambda_{270nm} + b_4\lambda_{310nm}$$

$$D.Q.O. = -0.01884 + 0.7081\lambda_{220nm} - 1.093\lambda_{260nm} + 1.315\lambda_{270nm} + 0.2462\lambda_{310nm}$$

Con el modelo matemático encontrado se establecieron los parámetros de calidad del mismo con los cuales se verifica la aptitud de predicción del modelo.

Figura 5: Ejemplo de algunos espectros de las muestras



Elaborado: Fernando Cárdenas Calle

2.2 Parámetros de Calidad del Modelo

La calidad del modelo en predicción viene representado en R^2_{Cv} , es decir el modelo matemático encontrado nos permite la predicción de valores de la Demanda Química de Oxígeno en muestras de efluentes de las industrias lácteas. El análisis de calidad del modelo se basaba en hallar el mayor R^2_{Cv} y el menor índice de correlación K entre las variables predictoras ya que esto es un requerimiento de la regresión multivariante de mínimos cuadrados (OLS) (Todeschini, 1998).

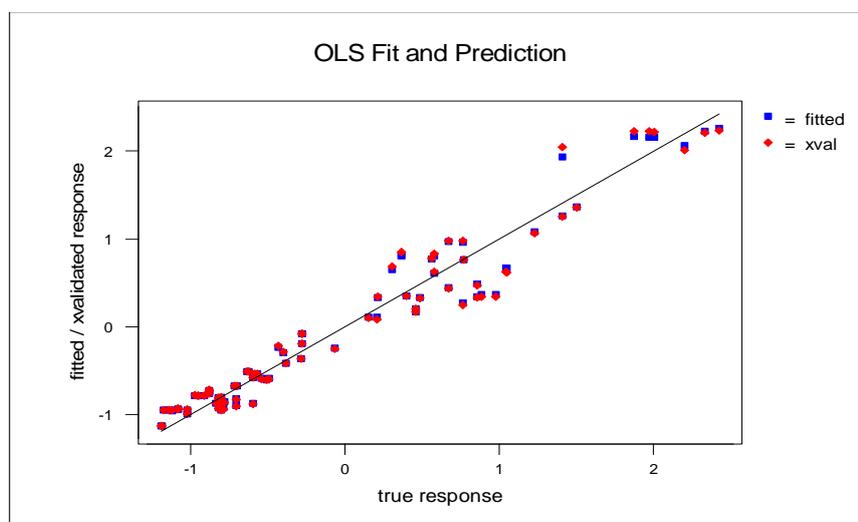
Los parámetros de calidad del modelo de regresión se presentan en la tabla No: 1

Tabla 1: Parámetros de calidad del modelo matemático

Parámetro	Valor
R^2	0.95
R^2_{CvLOO}	0.94
$R^2_{Cv(10\%)}$	0.939
K	0.569

En el gráfico se presenta la curva de calibración encontrada para el modelo donde se encuentran los valores calculados y los valores en predicción.

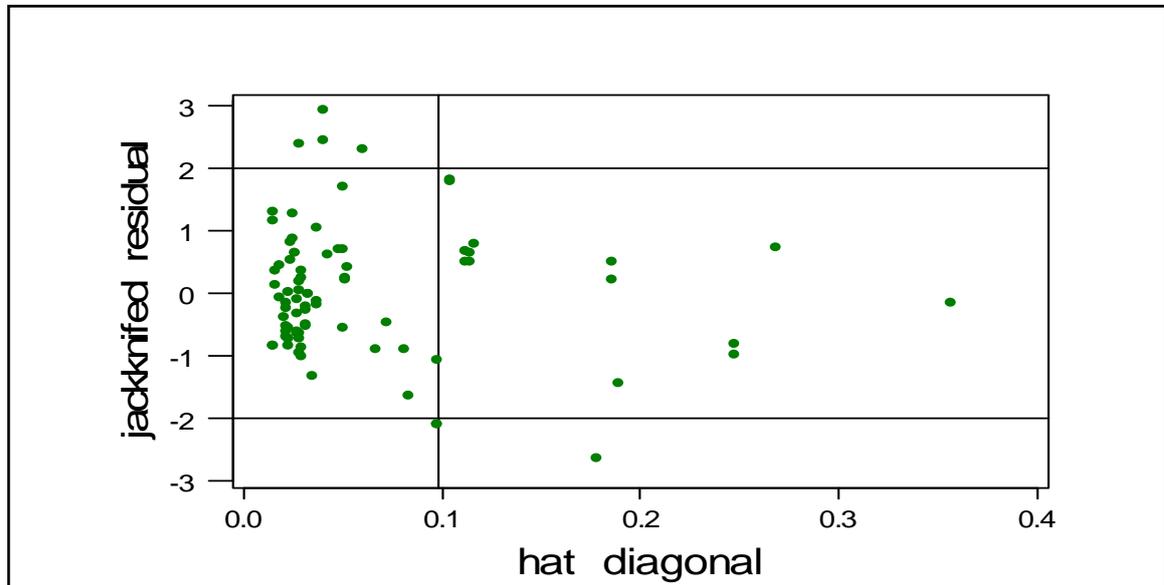
Figura 6: Relación entre DQO experimental y los valores calculados en interpolación y predicción por el modelo propuesto



Elaborado: Fernando Cárdenas Calle

Las características generales del modelo matemático encontrado y la distribución de los datos experimentales se representan en el gráfico de Williams de la figura No: 7.

Figura 7: Gráfico de Williams del modelo propuesto



Elaborado: Fernando Cárdenas Calle

La inspección del gráfico de la figura No: 3 muestra que:

- a) 6 valores calculados sobre un total de 81 (7.4%) presentan residuos anómalos, es decir que tienen una probabilidad inferior al 5% de pertenecer a la población problema.
- b) 13 puntos experimentales sobre 81 (16%) presentan un “leverage” que supera el nivel crítico de aproximadamente 0.1.

Con lo que se puede decir que el modelo matemático propuesto sirve para encontrar datos en predicción los cuales servirán para la determinación de la Demanda Química de Oxígeno de forma más eficiente y sobre todo sin afectar al medio ambiente al no introducir sustancias nocivas en el agua de análisis.

2.3 Validación Interna del Método

Cuando se desarrolla un método nuevo mediante procedimientos aceptables de investigación o se modifica un método existente con el fin de reunir requisitos

especiales, se precisa de un proceso de validación en tres etapas: determinación de la exactitud y la precisión de un operador único, análisis de muestras desconocidas preparadas de forma independiente y la determinación de la rigurosidad del método (APHA, 1992).

Para realizar el procedimiento de la validación interna de un método se debe realizar las operaciones de laboratorio por un analista único el cual debe determinar los siguientes parámetros: determinar los diferentes límites de cuantificación, la exactitud, y la precisión. Para realizar estas determinaciones se analizan como mínimo siete muestras. Se recomienda el empleo de varias concentraciones de la sustancia analizada para determinar la precisión y la exactitud la cual revelará la naturaleza de las relaciones existentes entre las características del método y la concentración de la sustancia analizada (APHA, 1992).

Para realizar el cálculo de los parámetros indicados anteriormente se determinaron 10 blancos que son muestras testigo. Estas muestras tuvieron el mismo tratamiento que las muestras que fueron utilizadas para la construcción de la curva de calibración. Con estos valores se calculó el valor que proporcionaban los valores del cero en la curva y con esos valores se realizó el cálculo de la exactitud y la precisión. Para la determinación de los diferentes límites de cuantificación se utilizó el coeficiente más bajo encontrado en el modelo matemático de la DQO que fue el coeficiente de 310nm que es 0.2462, el cual tiene la menor pendiente y por consiguiente la menor precisión (Massart et al, 1997).

Para determinar los diferentes límites de cuantificación del método se aplicaron las siguientes fórmulas (Massart et al, 1997):

:

$$LC = \mu_{(bl)} + 3\sigma_{(bl)}$$

$$LD = LC + 6\sigma_{(bl)}$$

$$LQ = LD + 10\sigma_{(bl)}$$

Para obtener la incertidumbre (μ) del método, se debe calcular el valor promedio de la diferencia entre los valores reales y los valores encontrados de las muestras testigo (blanco). Luego de realizar este cálculo se suma al valor la desviación estándar encontrada. Para cada límite cambia el valor de la desviación estándar, para el límite crítico tres valores de la desviación estándar, para el límite de detección por seis valores y para el límite de cuantificación por diez. Luego de estimar estos valores se los divide para el coeficiente más bajo encontrado en el modelo que es 310nm (0.2462) y se obtienen los valores de los límites. Estos valores son presentados en la tabla 2 que indican todos los parámetros obtenidos de la curva de calibración de la Demanda Química de Oxígeno (Massart et al, 1997).

Tabla 2: Parámetros encontrados en la validación interna del método

Parámetro	ppm
Exactitud	2
Precisión	6
Límite Crítico	15
Límite de Detección	22
Límite de Cuantificación	32

Los parámetros estudiados del método tienen mejores características que el método clásico de análisis de la Demanda Química de Oxígeno, por lo cual puede ser utilizado para la determinación de la DQO en las aguas de efluentes de de las industrias lácteas.

2.4 Validación Externa del Método

Este proceso se llevó a cabo para verificar la capacidad predictiva del modelo matemático encontrado frente a un conjunto de datos externos, esta validación tiene una mayor rigurosidad que los métodos aplicados anteriormente para la evaluación del modelo matemático. Para este proceso se realizó la valoración de un test externo el cual fue construido con 12 muestras independientes del modelo matemático encontrado para la determinación de la DQO.

En las 12 muestras independientes se determinó el DQO clásico y la absorbancia del espectro de bandas luego de esto se construyó un modelo matemático con las bandas ya seleccionadas por el algoritmo genético que para este propósito son: 220, 260, 270 y 310nm luego de esto se procedió a la validación externa del método mediante la aplicación de la siguiente fórmula (Todeschini, 1998):

$$Q^2 = 1 - \frac{P R E S S}{T S S}$$

Donde TSS es la suma total de los cuadrados y se calcula mediante la fórmula:

$$T S S = \sum_i (y_i - \bar{y})^2$$

Donde esta fórmula está referida al modelo matemático encontrado para la determinación de la DQO.

PRESS es el error predictivo de la suma de cuadrados y se determina de la siguiente manera (Todeschini, 1998):

$$P R E S S = \sum_i (y_i - \hat{y}_{i/i})^2$$

Esta expresión está referida al test de evaluación externa, al obtener estos dos valores se obtiene el porcentaje de varianza explicada del modelo en predicción el cual para este modelo matemático fue de 0.7%. El cual es un valor alto para una validación externa que fue construida con pocos datos, este valor garantiza que el método es confiable y sirve para la estimación de la Demanda Química de Oxígeno en muestras de efluentes de industrias lácteas.

2.5 Monitoreo de la Planta de Lácteos

Con el método propuesto es factible llevar a cabo un monitoreo de la Planta Piloto de Lácteos debido al costo del análisis y su fácil procesamiento en el laboratorio. El monitoreo determinó que el proceso más contaminante fue la producción de quesos, donde estas aguas residuales son las que más sólidos en suspensión presentan debido al cuajo. Además se desecha en su totalidad el suero de la leche lo cual es un proceso

deficiente. El resto de manufactura de productos presenta vertidos menos contaminantes al ser más eficientes en relación a los productos de desecho (Sans, 1999).

Figura 8: Marmita de capacidad de 100L para la elaboración de quesos



Elaborado: Fernando Cárdenas Calle

2.6 Costo del Método

Una determinación normal de la Demanda Química de Oxígeno se encuentra entre 6 y 8 dólares, el costo del análisis está determinado por el costo de las sustancias químicas, el valor de la energía eléctrica para la digestión que tiene una duración de dos horas, el costo del agua que para el análisis es de 6 litros, el costo del papel para la limpieza y el costo de operación del analista. Se realizó una estimación del costo del método el cual fue de 50 centavos ya que este método no utiliza sustancias químicas que elevan el costo del análisis, no se realiza la digestión de la muestra y se utiliza poca cantidad de papel para la limpieza.

2.7 Beneficios al Medio Ambiente

El método propuesto nos sirve para la estimación de la Demanda Química de Oxígeno en efluentes de industrias lácteas, con lo cual este método tiene muchas ventajas competitivas sobre el método clásico de análisis de la Demanda Química de Oxígeno, sobre todo en el aspecto Medio Ambiental ya que al utilizar un modelo matemático para la estimación de la Demanda Química de Oxígeno no se introduce sustancias químicas peligrosas al agua de análisis.

No se incorporan metales pesados como el Cromo el cual al llegar a las aguas superficiales afectaría a los peces debido a que el metal daña las agallas de los peces y afecta su respiración. Si es ingerido por animales puede causar problemas respiratorios, defectos de nacimiento, infertilidad y formación de tumores. Si llega por el alcantarillado a las plantas de tratamiento el Cromo produce problemas debido a que su proceso de descontaminación es costoso, si no se lo trata. El Cromo llega a los lodos siendo un problema ambiental grave ya que a los mismos se les debe dar una disposición correcta para que no afecten al medio ambiente (Sawyer, 2001).

El pH es un factor importante para el desarrollo de los organismos, si se afecta el pH como es en el caso de los residuos de la determinación de la demanda química de oxígeno se desechan por la alcantarilla y llegan a aguas superficiales, los organismos al ser expuestos a un pH ácido entran en estrés y luego se produce su muerte. Como en la determinación de la demanda química de oxígeno se utiliza el ión plata como catalizador, este elemento también forma parte de los desechos de la determinación el cual si llega a aguas superficiales afectaría al medio incrementando el pH y matando a los animales que habitan y además aumenta la dureza del agua alterando al medio receptor (Sawyer, 2001).

Si se utiliza el método de reflujo cerrado para la determinación de la demanda química de oxígeno esta se determina mediante una titulación mediante sulfato ferroso amónico la cual elimina hierro (III), el cual es persistente en el medio ambiente y puede afectar a los organismos que estén expuestos a este metal. Además afecta a los macroinvertebrados debido a que se tapan las agallas de los organismos disminuyendo su capacidad de respiración (Glynn, 1999).

Con el método propuesto se elimina del análisis todas estas sustancias peligrosas para el medio ambiente, gracias a la construcción de un modelo matemático por medio del cual se puede estimar la demanda química de oxígeno. Se debe tener en cuenta que para la construcción de la matriz de datos en la primera fase del método si se deben utilizar sustancias peligrosas pero estas deben tener una disposición adecuada para que no afecten al medio ambiente (Peralta, 2002).

DISCUSION

Las herramientas analíticas convencionales del análisis de aguas tiene altos niveles de confianza pero no son ambientalmente amigables, emplean en sus protocolos sustancias tóxicas que pueden afectar al analista, y al medio receptor que recibe las aguas de desecho de los laboratorios; lo cual está estimulando el desarrollo de las llamadas técnicas de análisis limpias. Estas técnicas en los últimos años han tomado fuerza en la implementación de la llamada química verde la cual promulga la mínima utilización de sustancias que puedan afectar a los seres humanos y al medio ambiente sin perder la efectividad del análisis. Estas herramientas están pasando a formar parte de las rutinas de los laboratorios, ya que estos métodos están siendo investigados para su posterior aplicación. Estas técnicas al no utilizar sustancias químicas o minimizar su cantidad son metodologías que tienen menor costo, por lo cual el análisis propuesto en este trabajo investigativo tiene un costo menor que el método clásico de análisis de la DQO (Rocha, 2001).

Estas técnicas se basan en no producir residuos como es el caso de encontrar un modelo matemático predictivo el cual pueda estimar un compuesto de interés ambiental como es el caso del método propuesto en el presente estudio. Las aguas residuales no deben ser tóxicas, esto se logra al no introducir sustancias peligrosas como es el caso del análisis de nitratos en donde se introduce cadmio, por esto se han desarrollado métodos que tengan los mismos resultados que los métodos clásicos. Otro caso es la determinación de amonio desarrollado por Rocha *et al.*, (2001); en donde el fenol es reemplazado por el ácido salicílico, sin perder la validez analítica del método.

Existen procesos de minimización de compuestos como es el proceso de inyección para el análisis de fósforo mediante el método del molibdato, en estos métodos se minimiza el consumo de agentes al minimizar una alícuota de muestra y reactivos presentes en el análisis, esto consiste en una simulación del análisis convencional mediante un sistema automatizado que transporta las sustancias por medio de una bomba peristáltica que empuja la muestra y los agentes por medio de un módulo analítico en donde se da la

reacción de color que es percibida por un registrador que para el caso es un UV-vis; en estos métodos de inyección se reducen los residuos a un 90% (Rocha *et al.*, 2001). Otros métodos han sido desarrollados de los procesos convencionales siguiendo el mismo procedimiento patrón minimizando al máximo las sustancias sin perder los resultados de los métodos clásicos de análisis.

Luego del análisis de una muestra de importancia ambiental, es fundamental el reciclaje del agua y su reutilización en el sistema, esta es una alternativa viable considerando la parte ambiental y económica. En trabajos de investigación sobre este tema se indica que, para realizar este proceso, se debe estudiar el desecho ya que es muy diverso y mediante el estudio se pueden dar las posibles formas de reciclaje del mismo. Si el residuo no puede ser reciclado se tiene que implementar un sistema de tratamiento en línea el cual destruya o disminuya el poder contaminante de las sustancias utilizadas en el análisis, en este proceso se puede llevar a cabo procesos químicos, físicos, o descomposición microbiana (Rocha *et al.*, 2001).

El reemplazo de las viejas técnicas por metodologías nuevas que tienen que ser probadas resulta de interés fundamental para el desarrollo de la Química Analítica y para el cuidado del Medio Ambiente. El estudio realizado se orientó a la minimización de compuestos tóxicos en la primera fase del estudio, en donde se realiza la determinación de la DQO mediante el método clásico utilizando una metodología semi-micro, luego se construye un modelo matemático el cual predice la DQO en efluentes de Industrias Lácteas, en esta fase no se introduce ninguna sustancias tóxicas al Medio Ambiente.

Rocha resalta que al realizar un procedimiento verde se debe estudiar y comprobar que las características que presenta el método estudiado sean las mismas que el método clásico de análisis. Siguiendo esta premisa se presentan los resultados hallados en el modelo matemático encontrado (Tabla 1) los cuales nos indican que posee una excelente capacidad de predicción la cual fue estimada al realizar una validación más robusta del modelo matemático al encontrar un R^2_{Cv} (10%) en donde se sacan al azar el 10% de los datos y se reproduce el modelo, se observó que el modelo se encuentra estable al realizar la perturbación lo cual nos confirma la capacidad predictiva del modelo matemático encontrado para la determinación de la Demanda Química de

Oxígeno en efluentes de industrias lácteas. Además se obtiene un valor elevado de correlación (0.95) entre las absorbancias y la Demanda Química de Oxígeno, entonces mediante este método se pueden predecir valores desconocidos de la DQO.

En el presente estudio se validan los resultados del método matemático de dos maneras mediante una validación LMO y se realiza una validación con un test externo que es una validación más fuerte, mientras en el análisis de Peralta no se realiza ninguna validación en su análisis, solo muestra los resultados obtenidos en la regresión PLS. En la evaluación del método espectrofotométrico para la determinación de la DQO en una industria de papel se determinaron nueve variables las cuales retienen el total de la información de los datos (Peralta, 2005). Los algoritmos genéticos determinaron cinco variables las cuales son las más importantes en el modelo matemático. Esta propiedad es una ventaja metodológica en relación al método propuesto por Peralta ya que al momento de leer las absorbancias se ahorra tiempo en el momento del análisis.

Una deficiencia del método de Peralta es la correlación entre las variables que en su análisis es de 90%, mientras que en el presente estudio es del 50% lo cual facilita el cálculo de mínimos cuadrados (OLS) donde la mínima correlación entre las variables es una condición para el desarrollo del método. Las características del modelo matemático se determinan mediante el gráfico de Williams (Figura 3) este determinó que se encuentran once “leverage”, estos son puntos predictivos demasiado altos encontrados en el modelo, Peralta determina la misma cantidad de “leverage”, mientras que el número de “outliers” son menores en este estudio. Al revisar las características que se encuentran en el presente análisis y el estudio de Peralta se determinó que el método sirve para la estimación de la Demanda Química de Oxígeno en muestras de importancia ambiental ya que los modelos matemáticos encontrados tiene buenas características de predicción.

Las características del método propuesto (Tabla 2) con el método clásico de análisis de la DQO se determinó que: la precisión del método clásico es de 13 ppm y la exactitud es de 6.5 ppm, con lo que se puede afirmar que el método propuesto tiene mejores características que el método clásico ya que la exactitud calculada para el método es de 2 ppm y la precisión tiene un valor de 6 ppm. Los límites de cuantificación tienen los mismos valores que son encontrados en el método clásico de análisis, lo cual permite

evaluar muestras de industrias lácteas con la misma calidad que se encuentra con el método clásico.

CONCLUSIONES

Con el método propuesto para la determinación de la DQO se reducen significativamente las sustancias químicas perjudiciales para el Medio Ambiente ya que para construir la matriz de datos se necesita realizar el método clásico de análisis de la DQO que se realizó mediante el sistema semi-micro con lo cual se reduce la cantidad de sustancias peligrosas al Medio Ambiente. Luego de construir el modelo matemático ya no se introducen sustancias tóxicas ya que el modelo matemático encontrado es un modelo en predicción el cual nos sirve para estimar la DQO. Sería de gran importancia el estudio de la disposición adecuada de los residuos del análisis ya que al realizar este estudio el método es más amigable con el medio ambiente.

La construcción del modelo matemático se facilitó al utilizar técnicas de análisis multivariante que ayudan a encontrar toda la información de un conjunto de datos. Luego del análisis se posibilita la toma de decisiones más acertada para solucionar un problema planteado. Para obtener un modelo matemático de buenas características se optó primero por seleccionar las variables importantes para el modelo mediante la utilización de los Algoritmos Genéticos, estos son métodos muy importantes que están teniendo un gran auge en el análisis de datos, gracias a estas herramientas que pueden ser aplicadas en el campo químico biológico se consigue la implementación de innovaciones tecnológicas que son útiles para la solución de problemas medio ambientales.

Además la selección de las variables tiene una ventaja metodológica en el análisis propuesto, ya que al saber cuales son las variables importantes solo se tienen que calcular las absorbancias de las longitudes de onda seleccionadas por el algoritmo genético, esto hace que el método tenga menor tiempo de análisis lo cual es una ventaja competitiva con el método clásico de análisis de la DQO. El modelo matemático encontrado tiene excelentes características en predicción ya que al perturbar el modelo con un 10% de objetos quitados al azar posee un $Q^2_{(10\%)} = 0.93$, lo cual es de gran

importancia ya que el modelo nos sirve para la estimación de la DQO en efluentes de industrias lácteas.

Con el método propuesto el costo del análisis baja debido a que no se utilizan sustancias químicas, la digestión no se realiza y se ahorra energía eléctrica, no se consume agua. En el método propuesto se necesita en mínima cantidad el papel y se debe añadir el costo de los servicios prestados por el analista con lo cual se ha determinado que el valor del análisis propuesto es de cincuenta centavos por muestra analizada; hay que tener en cuenta que se debe construir el modelo matemático y tener un espectro de UV-vis lo cual se transforma en una inversión grande pero el costo beneficio del método a largo plazo es importante.

Al reducir el costo del análisis se puede realizar un monitoreo de la planta en donde se lograría ver que procesos son los más contaminantes o se observaría en que procesos la planta está siendo deficiente. Se conseguiría optimizar esos procesos mediante la medición de los desechos debido a que una industria es más eficiente mientras menos residuos emiten al Medio Ambiente.

Se pueden implementar nuevas técnicas de análisis que sean amigables con el Medio Ambiente y que tengan los mismos resultados de los análisis clásicos, con el estudio de los métodos y la ayuda de las herramientas del análisis multivariante de datos se puede llegar a establecer nuevos protocolos de análisis que tengan las mismas características que los clásicos y que no afecten al Medio Ambiente.

BIBLIOGRAFIA

ALVAREZ, F. 2006. El Sector Lácteo en el País. Disponible en World Wide Web: www.hoy.com.ec

APHA, WPCF, AWWA, 1998. Métodos Normalizados para el Análisis de Aguas Potables. American Public Health Association. Madrid, España. 5-12 p.

BIANUCCI, G. Rivaldone, E. 1992., Il Trattamento delle Acque Inquinata. Ulrico Hoepli Editore, Milano, Italia. 287:318 p.

DELGADO, J. 1992. Elementos de Redes Neuronales y Algoritmos Genéticos. Neuronales LTDA. Colombia, Bogotá. 81: 192p.

DRAPER, D., Smith, H., 1966. Applied Regression Analysis. John, Wily & Sons, Inc. New York, USA. 44:81p.

FRANK, I., Todeschini, R. 1994. The data analysis handbook. Elsevier Science. Amsterdam, The Netherlands. 333-335 p.

FUNDACION NATURA. 1999. Potencial Impacto Ambiental de las Industrias en el Ecuador. EDUNAT. Quito, Ecuador. 25:36p.

GLYNN, H. Gary, W. 1999., Ingeniería Ambiental. Segunda Edición. Prentice Hall. DF, Mexico. 384:392 p.

LEARDI, R. 2001. Genetic algorithms in chemometrics and chemistry: a review. J. Chemometrics, 15: 559-569.

MASSART, D. Vandeginste, B. Buydens, L. Jong, S. Lewi, P. Verbeke, J. 1997. Handbook of Chemometrics and Qualimetrics. Part A. Elsevier Science B. V. Ámsterdam, The Netherlands. 805:844 p.

MELOUN, M. Forina, M. Militky, J. 1994. Chemometrics for Analytical Chemistry. Volume 2. Ellis Horwood. Great Britain. 273p.

PERALTA, P., 2005. Utilizacáo de regressáo multivariata para avaliacao espectrofotometrica de Demanda Química de Oxigenio en amostras de relevancia ambiental, 28(5): 383-841

ROCHA, F. Nabrega, J.; Fatibello, O. 2001. Flow analysis to greener analytical chemistry. The Royal Society of Chemistry. Green Chemistry: 3, 216 - 220 p.

SANS, F. Ribas, P., 1999. Ingeniería Ambiental Contaminación y Tratamientos. Alfaomega. Barcelona, España. 84 p.

SAWYER, C. 2001. Química para Ingeniería Ambiental. Mc Graw Hill. Cuarta Edición. Bogotá, Colombia. 589: 595 p.

SEOANES, M. 2000. Tratado de reciclado y recuperación de productos de los residuos. Alfa-omega. Madrid. 498 p.

R. TODESCHINI, D Ballabio, V. Consoni, A Mauri and M. Pavan. 2004. MobyDigs- Version 1.0-2004-Copyright Talete srl.

TODESCHINI, R. , 1998. Introduzione Alla Chemiometria. Edises. Napoli, Italia. 171:178p.

ANEXOS

Anexo 1: Fotografía de la Planta de Lácteos



Anexo 2: Fotografía de la Producción de Quesos en la Planta



Anexo 3: Fotografía Espectrofotometro UV-vis



Anexo 4: Fotografía del registrador de datos



Anexo 5: Fotografía Muestras y digestor Kjeldahl



Anexo 6: Matriz de Datos n x p utilizada para el calculo del modelo matemático

Muestra	200	210	220	230	240	250	260	270	280	290	300	310	COD
1	0.12729	-0.014	0.08903	0.034729	-0.05138	-0.0304	-0.05961	-0.15261	-0.2445	-0.40248	-0.45074	-0.3993	-0.2759
2	0.14723	-0.014	0.13616	0.142508	-0.05138	-0.0304	-0.05961	-0.09769	-0.1821	-0.40248	-0.45074	-0.3993	-0.2759
3	0.13726	-0.014	0.13616	0.106582	-0.05138	-0.0304	-0.05961	-0.09769	-0.2445	-0.40248	-0.45074	-0.3993	-0.2759
4	-0.5524	-0.627	-0.58586	-0.58246	-0.53305	-0.5126	-0.54274	-0.59951	-0.6797	-0.70548	-0.61371	-0.4544	-0.8805
5	-0.5441	-0.6191	-0.57599	-0.56742	-0.53305	-0.4947	-0.52192	-0.57651	-0.6536	-0.67819	-0.61371	-0.4445	-0.8805
6	-0.5483	-0.627	-0.58586	-0.6276	-0.53305	-0.5126	-0.52192	-0.59951	-0.6797	-0.70548	-0.63328	-0.4544	-0.8805
7	-0.8091	-0.8148	-0.82231	-0.76602	-0.54173	-0.526	-0.59521	-0.66637	-0.7397	-0.73459	-0.62258	-0.4528	-1.082
8	-0.7743	-0.8038	-0.81546	-0.80775	-0.64439	-0.6251	-0.66739	-0.73017	-0.8121	-0.81028	-0.66329	-0.4666	-1.082
9	-0.7859	-0.8093	-0.80862	-0.81818	-0.64439	-0.6251	-0.66739	-0.73017	-0.8121	-0.81028	-0.66329	-0.4666	-1.082
10	0.35045	0.36956	0.34632	0.348354	0.34135	0.33179	0.330044	0.31498	0.16363	-0.1471	-0.29791	-0.3058	0.4606
11	0.32403	0.34436	0.34632	0.348354	0.29453	0.21875	0.264205	0.242243	0.16363	-0.0608	-0.3598	-0.3372	0.4606
12	-0.4115	-0.4199	-0.41562	-0.44329	-0.27578	-0.2758	-0.31371	-0.39824	-0.4837	-0.52467	-0.50851	-0.3973	-0.4929
13	-0.4115	-0.394	-0.41562	-0.46774	-0.25172	-0.2177	-0.27988	-0.36087	-0.4837	-0.52467	-0.50851	-0.3973	-0.5937
14	-0.6493	-0.6022	-0.63734	-0.63251	-0.35215	-0.3743	-0.4314	-0.53151	-0.6167	-0.64615	-0.56521	-0.4268	-0.803
15	-0.7237	-0.6968	-0.72938	-0.7218	-0.50279	-0.4955	-0.57262	-0.64852	-0.7718	-0.76184	-0.64818	-0.4604	-0.5937
16	-0.8249	-0.7903	-0.8285	-0.828	-0.58219	-0.5833	-0.6621	-0.72969	-0.8238	-0.79735	-0.66766	-0.4665	-1.1208
17	-0.8385	-0.805	-0.84222	-0.84194	-0.58219	-0.5833	-0.6621	-0.75101	-0.848	-0.82264	-0.66766	-0.4665	-1.02
18	0.61457	0.84825	0.81441	0.87171	1.13731	1.12306	1.120109	1.115083	0.99003	0.629615	0.13532	-0.1488	0.7706
19	0.58816	0.84825	0.81441	1.014444	1.23095	1.12306	1.120109	1.115083	0.99003	0.629615	0.13532	-0.1175	0.5613
20	-0.2486	-0.1998	-0.22319	-0.22324	0.01295	0.01465	-0.07688	-0.13659	-0.2289	-0.30292	-0.4131	-0.3651	-0.2836
21	-0.269	-0.2386	-0.25526	-0.27214	-0.05923	-0.1015	-0.14454	-0.21135	-0.3138	-0.39162	-0.4449	-0.3812	-0.3844
22	-0.7151	-0.6881	-0.72453	-0.7269	-0.50094	-0.5124	-0.57299	-0.66658	-0.768	-0.75575	-0.63473	-0.4696	-0.7022
23	-0.6482	-0.5998	-0.70023	-0.61572	-0.35506	-0.3583	-0.44477	-0.52494	-0.6071	-0.6213	-0.56241	-0.4207	-0.7022
24	-0.7991	-0.7617	-0.80138	-0.79461	-0.52535	-0.5477	-0.63019	-0.71038	-0.7958	-0.79117	-0.65716	-0.4673	-1.02
25	-0.7923	-0.7488	-0.79073	-0.79461	-0.52535	-0.5284	-0.6077	-0.69796	-0.7816	-0.7617	-0.64659	-0.4566	-0.803
26	0.5533	1.38237	1.35864	1.385551	0.83391	0.87437	0.988991	0.800861	0.71897	0.35345	0.00164	-0.2154	0.7659
27	0.61193	0.93492	1.22009	1.10389	0.83391	0.79072	0.801449	0.800861	0.71897	0.35345	-0.08996	-0.2154	0.6736
28	0.65102	1.38237	1.45101	1.174305	1.38827	1.37627	1.483539	1.446763	1.45282	0.864355	0.18483	-0.1225	1.4114
29	0.57284	1.38237	1.45101	1.455966	1.38827	1.29262	1.386097	1.446763	1.33051	0.864355	0.32727	-0.1225	1.5036
30	0.65102	2.05355	2.23616	2.864269	2.77418	2.71465	2.847718	2.846217	2.67589	2.141619	1.10081	0.24912	2.4258
31	0.65102	2.16541	2.25925	2.582609	2.91277	2.71465	2.750277	2.738567	2.7982	2.013892	0.91761	0.20267	2.3336
32	0.77032	1.01592	0.82194	1.06318	2.39664	2.5185	2.504993	2.419328	2.38066	2.159236	1.16398	0.30679	1.227
33	1.47799	1.85168	1.53861	1.956157	3.47191	3.38384	3.345005	3.347348	3.22416	2.81989	1.47984	0.46696	1.4114
34	1.94977	2.17313	1.77749	2.077562	3.47191	3.67228	3.513007	3.532952	3.64591	3.040108	1.47984	0.46696	1.8725
35	0.43333	0.5659	0.38399	0.620696	1.67979	1.65317	1.664981	1.676912	1.53717	1.278365	0.53227	0.06652	0.3048
36	0.43333	0.5659	0.34417	0.499291	1.44084	1.50895	1.328977	1.305704	1.32629	1.058147	0.37435	-0.0136	0.8581
37	0.50073	0.5659	0.50343	0.559994	1.20189	1.22051	1.160974	0.934496	0.90454	0.617711	0.21642	-0.0937	0.7659
38	0.39963	0.37303	0.34417	0.438588	0.84346	1.07628	0.992972	0.934496	0.69367	0.617711	0.21642	-0.0937	0.8581
39	0.39963	0.37303	0.4238	0.499291	0.96294	1.07628	1.160974	1.1201	1.11542	0.837929	0.37435	-0.1737	0.6736
40	-0.1465	-0.1632	-0.12697	-0.05847	0.30063	0.4137	0.380615	0.248568	0.22531	0.233128	-0.06739	-0.1996	-0.0641
41	0.47313	0.03948	0.03948	0.834476	0.30063	0.4137	0.380615	0.248568	0.11453	0.00174	-0.15036	-0.2416	0.3907
42	0.45543	0.07325	0.70971	0.834476	0.30063	0.4137	0.380615	0.248568	0.11453	0.00174	-0.15036	-0.2416	0.2126
43	0.43772	0.03948	0.70971	0.834476	0.42617	0.4137	0.380615	0.248568	0.11453	0.00174	-0.15036	-0.2416	0.4892
44	2.85964	2.04455	2.17564	1.956157	-0.11233	-0.2217	-0.01504	0.563288	0.06104	-0.26316	-0.41529	-0.4709	1.9721
45	1.03991	0.63019	0.66269	0.499291	-0.35128	-0.5102	-0.51905	-0.17913	-0.3607	-0.48338	-0.54795	-0.4709	1.0419
46	1.02306	0.5659	0.58306	0.377886	-0.47076	-0.6544	-0.51905	-0.36473	-0.5716	-0.48338	-0.54795	-0.4709	0.8869
47	1.88237	1.33737	1.37935	0.984913	-0.23181	-0.3659	-0.18304	-0.17913	-0.1498	-0.48338	-0.57322	-0.4709	0.5768
48	-0.3417	-0.3341	-0.37249	-0.47195	-0.59023	-0.3659	-0.18304	-0.55034	-0.1498	-0.48338	-0.59375	-0.4709	0.8185
49	-0.2676	-0.4242	-0.43619	-0.47195	-0.68581	-0.6832	-0.68705	-0.62458	-0.7825	-0.79168	-0.66797	-0.4709	-0.6324
50	-0.1665	-0.1541	-0.30879	-0.42339	-0.63802	-0.6832	-0.72065	-0.73594	-0.7825	-0.83573	-0.54795	-0.4709	-0.5704
51	-0.6316	-0.707	-0.73878	-0.76333	-0.7336	-0.7409	-0.78785	-0.77306	-0.8668	-0.83573	-0.54795	-0.4709	-0.8185
52	-0.4294	-0.5527	-0.59545	-0.61764	-0.68581	-0.7121	-0.72065	-0.77306	-0.8246	-0.83573	-0.54795	-0.4709	-0.9735
53	-0.4631	-0.5656	-0.59545	-0.6662	-0.70971	-0.7121	-0.72065	-0.77306	-0.8246	-0.83573	-0.69956	-0.4709	-0.9115
54	-0.9079	-0.8999	-0.94582	-0.95757	-0.78139	-0.7121	-0.72065	-0.8473	-0.8246	-0.83573	-0.59375	-0.4709	-1.1906
55	-0.8405	-0.797	-0.85026	-0.83617	-0.61413	-0.6255	-0.68705	-0.73594	-0.8246	-0.74764	-0.66797	-0.4621	-1.1751
56	2.85964	2.04455	2.17564	1.956157	-0.11233	-0.2217	-0.01504	0.563288	0.06104	-0.26316	-0.41529	-0.4709	2.0031
57	3.36512	3.39462	2.81267	2.198967	0.12662	-0.2217	-0.18304	0.006476	0.06104	-0.48338	-0.54795	-0.4709	2.2047
58	1.03991	0.63019	0.66269	0.499291	-0.35128	-0.5102	-0.51905	-0.17913	-0.3607	-0.48338	-0.54795	-0.4709	1.0497
59	1.02306	0.5659	0.58306	0.377886	-0.47076	-0.6544	-0.51905	-0.36473	-0.5716	-0.48338	-0.54795	-0.4709	0.9799
60	2.05087	1.40166	1.37935	1.227724	-0.11233	-0.3659	-0.18304	-0.17913	-0.1498	-0.48338	-0.54795	-0.4709	0.3675
61	1.88237	1.33737	1.37935	0.984913	-0.23181	-0.3659	-0.18304	-0.17913	-0.1498	-0.48338	-0.57322	-0.4709	0.3675
62	-0.3417	-0.3341	-0.37249	-0.47195	-0.59023	-0.3659	-0.18304	-0.55034	-0.1498	-0.48338	-0.59375	-0.4709	0.7875
63	-0.2676	-0.4242	-0.43619	-0.47195	-0.68581	-0.6832	-0.68705	-0.62458	-0.7825	-0.79168	-0.66797	-0.4709	-0.6247
64	-0.1665	-0.1541	-0.30879	-0.42339	-0.63802	-0.6832	-0.72065	-0.73594	-0.7825	-0.83573	-0.54795	-0.4709	-0.5859
65	-0.6316	-0.707	-0.73878	-0.76333	-0.7336	-0.7409	-0.78785	-0.77306	-0.8668	-0.83573	-0.54795	-0.4709	-0.8185
66	-0.6349	-0.7199	-0.75471	-0.78761	-0.7575	-0.7698	-0.78785	-0.81018	-0.909	-0.83573	-0.54795	-0.4709	-0.834
67	-0.4294	-0.5527	-0.59545	-0.61764	-0.68581	-0.7121	-0.72065	-0.77306	-0.8246	-0.83573	-0.54795	-0.4709	-0.9503
68	-0.4631	-0.5656	-0.59545	-0.6662	-0.70971	-0.7121	-0.72065	-0.77306	-0.8246	-0.83573	-0.69956	-0.4709	-0.9503
69	-0.9079	-0.8999	-0.94582	-0.95757	-0.78139	-0.7121	-0.72065	-0.8473	-0.8246	-0.83573	-0.59375	-0.4709	-1.1828
70	-0.8405	-0.797	-0.85026	-0.83617	-0.61413	-0.6255	-0.68705	-0.73594	-0.8246	-0.74764	-0.66797	-0.4621	-1.144
72	-1.0381	-0.9983	-1.00922	-0.91648	-0.50936	-0.4125	-0.34911	-0.05206	0.77072	2.613224	3.81351	4.16883	0.5768
73	-1.0414	-1.0184	-1.04183	-0.96355	-0.60198	-0.5243	-0.49564	-0.30388	0.36007	1.588836	2.34714	2.93259	0.2065
74	-1.0414	-1.0184	-1.04183	-0.97797	-0.60198	-0.5243	-0.49564	-0.30388	0.36007	1.588836	2.34714	2.93259	0.1505
75	-1.0397	-1.0136	-1.04342	-0.9744	-0.62335	-0.5294	-0.49327	-0.36198	0.09654	0.846651	1.3094	1.62148	