DIFERENCIAS ESTADÍSTICAS ENTRE LOS MODELOS DE DISTRIBUCIONES DE POROS A PARTIR DE LA CURVA DE RECHAZO DE LAS MEMBRANAS DE ULTRAFILTRACIÓN

JOSÉ A. PITTOL VERA¹, ISABEL LLATAS^{1,2} Y SYLVANA DERJANI-BAYEH³

¹Departamento de Procesos y Sistemas, Universidad Simón Bolívar, Caracas-Venezuela, jpittol@usb.ve ²Centro de Estadística y Software Matemático, Universidad Simón Bolívar, Caracas-Venezuela, llatas@usb.ve ³Grupo TADiP, Departamento de Termodinámica y Fenómenos de Transporte, Universidad Simón Bolívar, Caracas-Venezuela, sderjani@usb.ve

Recibido: marzo de 2010

Recibido en forma final revisado: marzo de 2011

RESUMEN

La distribución efectiva de poros de las membranas de ultrafiltración, se obtiene a partir de datos experimentales de curvas de rechazo de diferentes tipos de membranas y solutos. Para ello, se adopta un modelo hydrodinámico de transporte y se comparan diferentes modelos estándares de densidad de probabilidad para describir la distribución efectiva de poros. Los modelos probabilísticos escogidos son Gamma, Log-normal y Weibull, por ser sesgados a la derecha. Con el algoritmo robusto de Levenberg-Marquardt para regresión no lineal, implementado en la subrutina *lsqcurvefit* de Matlab[®], se obtienen los parámetros optimizados de dichos modelos probabilísticos. Los intervalos de 95% confianza para los parámetros optimizados, se obtienen de la subrutina *nlparci* de Matlab. El análisis de varianza (ANOVA) se usa para probar el efecto del modelo de densidad de probabilidad en la predicción de los momentos centrales que caracterizan la distribución efectiva de poros, el coeficiente promedio de filtrado y el flujo promedio de agua. Con diferencias estadísticamente significativas de un 5 a 10%, el análisis (ANOVA) de dos vías y el de una vía indicaron que existe una diferencia relacionada con el modelo probabilístico escogido para predecir del tamaño promedio efectivo de poros y el coeficientes promedio de filtrado.

Palabras clave: Ultrafiltración, Curva de rechazo, Distribución efectiva de poros, Tamaño promedio efectivo de poro, Coeficiente promedio de filtrado.

STATISTICAL DIFFERENCES AMONG PORE DISTRIBUTION MODELS OBTAINED FROM REJECTION PROFILE OF ULTRAFITRATION MEMBRANES

ABSTRACT

The effective pore size distribution of ultrafiltration membranes' is obtained from the experimental rejection profile of different types of membranes and solutes. For this, the hydrodynamic transport model is assumed and different standards probability density functions are compared to describe the effective pore size distribution. The chosen probabilistic models are the gamma, log-normal and Weibull, because they are skewed to the right. The robust Levenberg-Marquardt algorithm for non-linear regression as implemented in Matlab[®] *lsqcurvefit* subroutine is used to provide the optimized parameters. The 95% confidence intervals for the optimized parameters are obtained from *nlparci* subroutine in Matlab[®]. The analysis of variance (ANOVA) is used to test the effects of the probability distribution model in the prediction of the central moments, the average sieving coefficient and the average water flux. With statistical significant differences between 5 and 10%, the two-way and one-way (ANOVA) analysis, indicated that the chosen probabilistic model introduced a significant difference in the predictions of the effective average pore size, and the average sieving coefficient.

Keywords: Ultrafiltration, Rejection profile, Effective pores size distribution, Average effective pore size, Average sieving coefficients.

INTRODUCCIÓN

El proceso de Ultrafiltración (UF) consiste en una separación física de baja presión (1-10 bar) que utiliza una barrera selectiva, llamada membrana, la cual es una pared delgada, de porosidad y tamaño de poros controlada, que logra separar partículas con un tamaño de poro entre 1-100nm, equivalente a pesos moleculares entre 0,5 y 500kdalton, aproximadamente. El proceso de ultrafiltración puede ser pensado como un mecanismo de tamizado, en el cual las moléculas se separan por diferencia de tamaños y formas, impulsadas a través del espesor de la membrana por la diferencia de presión, como se puede observar en el esquema de la figura 1. La selectividad resulta fundamental en el proceso de ultrafiltración, pues ésta es la que permite la recuperación eficiente de productos de alto valor en procesos tales como la separación de la proteína de suero de queso o la recuperación de una enzima en un sistema de reacción, entre otros (Wagner, 2001). La UF es una de las técnicas más competitivas porque tiene bajo costo energético y permite conservar la naturaleza de los productos tratados, punto muy importante cuando la meta consiste en la reutilización de productos.



Figura 1. Esquema general del ultrafiltrado.

Los parámetros indicativos del desempeño de una membrana de ultrafiltración, se fundamentan en el coeficiente promedio de filtrado o de tamizado y la permeabilidad promedia del solvente. Estos parámetros de transporte globales dependen de la distribución efectiva de poros de la estructura porosa de la membrana, siendo esta distribución efectiva de los tamaños de los poros una medida del área y el volumen disponible para el transporte que, a su vez, es responsable de la separación selectiva de los solutos.

Para predecir teóricamente el coeficiente promedio de filtrado y el flujo promedio de solvente de un proceso de ultrafiltración, se requiere suponer un modelo probabilístico para describir analíticamente la estructura porosa, acoplado a un supuesto y aceptado mecanismo de transporte.

Una de las técnicas de caracterización de la membrana de ultrafiltración que más se asemeja a las condiciones del proceso de ultrafiltración, se fundamenta en el desarrollo de la curva de rechazo, la cual es una curva operativa y una medida del coeficiente experimental de filtrado de la membrana como función del radio efectivo de los poros. Esta curva tiene una forma sigmoidea y representa la funcionalidad entre la retención o rechazo de soluto y el tamaño de soluto, y por lo tanto ésta es una medida indirecta de la distribución de tamaños de poros disponible para el transporte selectivo y de la permeabilidad del solvente (Michaels, 1980).

El objetivo principal de este trabajo de investigación, se

fundamenta en identificar diferencias estadísticas entre la predicción del radio promedio efectivo de la membrana, del coeficiente promedio de separación y el flujo promedio de solvente, a partir del modelo hidrodinámico de transporte, o del modelo de transporte restringido del soluto a través de una membrana de ultrafiltración, cuando diferentes modelos de densidad de probabilidad se usan para describir la distribución efectiva de los tamaños de poros de la membranas de ultrafiltración.

MARCO TEÓRICO

En el proceso de ultrafiltración, los solutos pueden ser suficientemente grandes para comportarse como partículas hidrodinámicas dentro del poro lleno de líquido. El transporte restringido se atribuye a la reducción de la movilidad del soluto y la restricción del espacio disponible dentro del poro. Las paredes del poro aumentan la fuerza de arrastre sobre las esferas que se trasladan en paralelo y disminuyen la velocidad de las esferas libremente suspendidas, en relación con la velocidad del fluido que se acerca. Razón por la cual ambos factores afectan el equilibrio de separación del soluto (Deen, 1987).

El enfoque de la teoría de transporte restringido o teoría hidrodinámica, consiste en el estudio del transporte de sólidos de tamaño finito en canales cilíndricos llenos de fluido. El transporte de las partículas depende del radio relativo de la partícula al radio de poro y de las interacciones hidrodinámicas con las paredes del canal. El transporte difusivo del sólido en el fluido posee características Brownianas (Deen, 1987). El solvente se puede considerar como un flujo continuo, debido a que el radio del soluto y del poro tienen órdenes de magnitud mayores que el radio molecular del solvente. En la figura 2, se muestra el esquema de transporte de una partícula de soluto atravesando un poro lleno de líquido.



Figura 2. Transporte de soluto a través de un poro.

Un modelo unidimensional para el transporte de soluto dentro del poro, puede derivarse haciendo un balance de las fuerzas de arrastre que impulsan el flujo de soluto en el solvente. Considerando que la concentración de soluto es radialmente constante y la resistencia de transferencia de masa en la entrada y salida del poro es insignificante, la ecuación del transporte promedio de soluto, $\langle N_s \rangle$, a través del poro presenta contribuciones difusivas y convectivas y viene dada por la ecuación (1):

$$\langle N_{S} \rangle = -D_{\infty}K_{d}\frac{d\langle C \rangle}{dz} + K_{c}\langle C \rangle \langle V \rangle$$
 (1)

donde: $\langle C \rangle$ es el promedio radial de concentración de soluto, $\langle V \rangle$ es el promedio radial de la velocidad de la solución en el poro que tiene un perfil parabólico de acuerdo con la ley de Hagen-Poiseuille, K_d es el inverso del coeficiente integral de arrastre, que es una medida del incremento de la fricción en el soluto, mientras que K_c es el coeficiente integral de retraso debido a los efectos de la pared (Deen, 1987). Los parámetros de resistencia convectiva K_c y difusiva K_d se determinan a partir de la resistencia que actúa en el soluto en la dirección axial, y de la suposición de trasporte en la línea central (Bungay y Brenner, 1973), y están definidos en las ecuaciones (2) y (3):

$$K_c = \frac{(2 - \phi) K_s}{2K_t}, \ 0 \le \lambda < 1$$
(2)

$$K_d = \frac{6\pi}{K_t}, \ 0 \le \lambda < 1 \tag{3}$$

donde: λ representa la relación del radio del soluto al radio del poro, ϕ es el coeficiente de partición, y las funciones hidrodinámicas $K_s y K_t$, son el resultado de aproximaciones asintóticas de la teoría del transporte restringido y tienen una precisión dentro del 10% para todos los valores de λ , calculados a partir de la ecuación (4):

$$\frac{K_t}{K_s} = \frac{9\pi^2 \sqrt{2}}{4(1-\lambda)^{5/2}} \left(1 + \sum_{n=1}^2 \left(\frac{a_n}{b_n}\right) (1-\lambda)^n\right) + \sum_{n=0}^4 \left(\frac{a_{n+3}}{b_{n+3}}\right) (4)$$

Los coeficientes a_n y b_n son presentados en la tabla 1.

Tabla 1. Coeficientes de la correlación de Bungay &Brenner, 1973 ecuación (4). (Deen, 1987).

a_n	b_n
-73/60	-7/60
-77,293/50.4	2,227/50,4
-22,5083	-3,9788
0,3363	-1,9215
-1,216	-4,392
1,647	5,006
-22,5083 0,3363 -1,216 1,647	-3,9788 -1,9215 -4,392 5,006

La condición de borde para resolver la ecuación diferencial (1) para el transporte de soluto en el poro es el coeficiente de distribución, ϕ , para sólidos esféricos y poros cilíndricos, viene dada por las restricciones geométricas $(1 - \lambda)^2$; esto supone que existe un equilibrio entre la concentración promedio del fluido en la entrada del poro $\langle C_0 \rangle$ y la concentración de bulto del fluido rechazado o retenido a la entrada del poro C_0 , que se expresa en la ecuación (5):

$$\varphi = \frac{\langle C_0 \rangle}{C_0} = \frac{\langle C_L \rangle}{C_L} = (1 - \lambda)^2$$
(5)

Similarmente se puede establecer la relación de concentración promedio del fluido a la salida del poro, $\langle C_L \rangle$, y la concentración de bulto del fluido en el permeado, C_L . El coeficiente de partición contabiliza la restricción de transporte debido a la disposición de las partículas relativo al poro. Si se suponen despreciables las interacciones estéricas entre el soluto y el poro, y entre soluto-soluto dentro del poro en comparación con los efectos de las paredes, entonces el flujo de soluto, $\langle N_s \rangle$, (ecuación (6)), a lo largo del poro resulta de la integración de la ecuación (1) y la aplicación de la condición de borde dada por la partición de equilibrio(φ), (5)

$$\langle N_s \rangle = \varphi K_c \langle V \rangle \frac{C_0 \exp(-\text{Pe}) - C_L}{\exp(-\text{Pe}) - 1}$$
 (6)

Donde: Pe, es número adimensional de Peclet y es la relación entre el transporte conectivo y el transporte difusivo definido en la ecuación (7):

$$Pe = \frac{L\langle V \rangle \varphi K_c}{D_{\infty} \varphi K_d} \tag{7}$$

La concentración del soluto en el permeado es determinada por la relación entre el flujo de soluto a lo largo del poro y el promedio radial de la velocidad de la solución en el poro, representada por la ecuación (8):

$$C_L = \frac{\langle N_S \rangle}{\langle V \rangle} \tag{8}$$

Combinando las ecuaciones obtenidas para el flujo de un soluto esférico en un poro cilíndrico (6) con la concentración de permeado (8) se puede obtener la expresión para el coeficiente de separación a través del poro, usando la ecuación (9):

$$S = \frac{C_L}{C_0} = \frac{\exp(Pe)\varphi K_c}{\varphi K_c + \exp(Pe) - 1}$$
(9)

Perfil de rechazo de una membrana de ultrafiltración (UF)

La selectividad de una membrana de ultrafiltración está determinada principalmente por la relación entre el diámetro hidrodinámico del soluto y del radio aparente del poro de la membrana. La curva de rechazo se obtiene experimentalmente a partir de la medición de la permeabilidad de solutos en óptimas condiciones de operación, donde se reduce al mínimo la polarización de concentración, la adsorción de solutos y la obstrucción de la membrana (Tkacik, 1991). Esto se logra con una solución de baja concentración, en un sistema operando a bajas presiones, y a alta velocidad de agitación o de alto flujo tangencial, para diferentes tamaños de solutos.

Las membranas se elaboran a partir de polímeros hidrofílicos, tales como polisulfona y acetato. Dextrano y polietilenglicol (PEG) son solutos que se usan para caracterizar la estructura porosa de las membranas poliméricas de ultrafiltración por ser solutos inertes, solubles en agua, y que no se adsorben a la superficie de la membrana. La curva de rechazo se obtiene usando un rango de solutos para todo el rango de poros de la membrana y tiene la forma de una curva sigmoidal que cubre todo el rango de separación. Los valores de rechazo o retención de soluto están entre uno (retención total) y cero (no retención).

El coeficiente de rechazo observado para un tamaño de soluto, R(PM), es una medida de la selectividad de la membrana y se define de acuerdo a la ecuación (10):

$$R(PM) = 1 - \frac{C_f(PM)}{C_r(PM)}$$
(10)

Donde: $C_f(PM)$ es la concentración de la solución permeada o filtrada y $C_r(PM)$ es la concentración de la solución retenida o rechazada. Y todas las concentraciones son función del peso molecular del soluto.

La capacidad de separación de la membrana se determina a través del *Peso Molecular de Corte*, PMC, que es la designación nominal del tamaño de poros en Daltons de las membranas de ultrafiltración. Este es un indicador de la separación por peso molecular de un soluto inerte a través de una membrana y se define como el peso molecular de un soluto inerte para el cual se obtiene un rechazo del 95% a través de la membrana. La distribución de peso molecular se determina usando cromatografía de gel de exclusión molecular (CGEM) o técnicas como HPLC. Para el dextrano la relación empírica de difusividad o de viscosidad intrínseca de uso común, que relaciona el tamaño aparente esférico del soluto con su peso molecular, se presenta en la ecuación (11), publicada por Granath para el rango de, 1000 $Da \le PM \le 300\ 000\ Da$ (Granath, 1958).

$$Log(r_{o}) = 0,470Log(PM) - 1,513$$
 (11)

Donde: r_s es el radio del soluto (en A^0). Para el polietilenglicol la relación entre el radio de soluto r_s y el peso molecular *PM* está basada en medidas de la viscosidad intrínsecas de Tam (Tam, 1991), que conducen a la ecuación (12), válida para el rango de peso molecular 200 $Da \le PM \le 40\ 000\ Da$.

$$r_s = 0.262 (PM)^{0.5} - 0.3 \tag{12}$$

Una vez que los datos son transformados, la curva de rechazo se obtiene para todo el rango de tamaños del solutos, y se grafica como una función del radio aparente del soluto. Así se genera una curva de forma sigmoidal en un rango de valores normalizados entre cero y uno. El rango de valores normalizado es una medida indirecta de la frecuencia de poros de un tamaño dado que están disponibles para el transporte. Los datos de la curva de rechazo se usan para obtener información sobre la distribución efectiva de poros de la estructura porosa de la membrana de ultrafiltración (UF).

Transporte global a través de la membrana de UF

Los parámetros de transporte globales tales como la permeabilidad del solvente y el coeficiente de separación o de rechazo, son indicativos del desempeño de una membrana de ultrafiltración. Estos parámetros se definen como el promedio de las propiedades microscópicas y dependen del supuesto modelo de distribución de los poros de la membrana y de un aceptado modelo de transporte. Dichos parámetros son coeficientes efectivos de transporte y son una medida de la selectividad y permeabilidad promedio de la membrana.

La permeabilidad promedio de una membrana limpia es una medida del radio efectivo medio. La disminución de la permeabilidad promedio es una medida del grado de obstrucción o de ensuciamiento de la membrana, durante el proceso de separación. Dicha permeabilidad, a través de los poros abiertos de la membrana, se determina teóricamente suponiendo que la estructura porosa tiene una distribución de poros cilíndricos y el líquido tiene características de un flujo tipo Poiseuille (ecuación (13)):

$$\langle J_{solvente} \rangle = \frac{\int Q(r) f(r) dr}{\int A(r) f(r) dr}$$
 (13)

Donde: $\langle J_{solvente} \rangle$ es la permeabilidad promedio del solvente a través de la membrana, Q(r) es el flujo volumétrico tipo Poiseuille a través de un poro cilíndrico, A(r) es el área transversal de un poro abierto disponible para el transporte y f(r) es la función de densidad de probabilidad o distribución de poros, que cuantifica la frecuencia de poros que están abiertos o que están disponibles para el transporte, de manera que la ecuación (13) se puede expresar en este caso como:

$$\langle J_{solvente} \rangle = \frac{\int \frac{\Delta P \pi r^4}{8 \mu L} f(r) dr}{\int \pi r^2 f(r) dr} = \frac{\Delta P}{8 \mu L} \frac{\int r^4 f(r) dr}{\int r^2 f(r) dr}$$
(14)

Donde: ΔP es la caída de presión a través del espesor, *L* el espesor de la membrana y μ la viscosidad de la solución.

El coeficiente promedio de filtrado o tamizado promedio $\langle S \rangle$ es una medida de la selectividad de la membrana y mide la fracción de soluto que se separa por diferencia de tamaño. Experimentalmente, se determina mediante el cociente de la concentración del soluto en la solución filtrada $\langle C_L \rangle$ y la concentración del soluto en la solución rechazada o retenida $\langle C_0 \rangle$ a partir de la ecuación (15):

$$\langle S \rangle = \frac{\langle C_L \rangle}{\langle C_0 \rangle} \tag{15}$$

Para condiciones de transporte convectivo, la relación entre el flujo del soluto y del solvente es una medida de la concentración del soluto en el filtrado.

$$\langle C_L \rangle \cong \frac{\langle J_{soluto} \rangle}{\langle J_{solvente} \rangle}$$
 (16)

Combinando las ecuaciones (15) y (16), el flujo del soluto dentro del poro queda descrito por la ecuación:

$$\langle J_{soluto} \rangle = \langle J_{solvente} \rangle \langle S \rangle \langle C_0 \rangle$$
 (17)

Con esto, para una distribución de poros, el flujo promedio de soluto a través de la membrana queda expresado como:

$$\langle J_{soluto} \rangle = \frac{\int J_{soluto} f(r) dr}{\int A(r) f(r) dr} = \langle C_0 \rangle \frac{\Delta P}{8 \,\mu L} \frac{\int_0^{max} f(r) dr}{\int_0^{rmax} f(r) dr}$$
(18)

El coeficiente promedio de filtrado se obtiene combinando las expresiones obtenidas por la teoría hidrodinámica para el coeficiente de filtrado, S, dentro un poro cilíndrico, el flujo promedio de soluto y el flujo promedio de solvente, ecuaciones (14), (15), (16) y (18)

$$\langle S \rangle = \frac{\int_{0}^{r_{max}} \frac{\exp(Pe)\varphi K_c}{\varphi K_c + \exp(Pe) - 1} r^4 f(r) dr}{\int_{0}^{r_{max}} r^4 f(r) dr}$$
(19)

El coeficiente promedio teórico de rechazo para la estructura porosa de la membrana es, por definición, uno menos el coeficiente promedio de filtrado.

$$Rec = 1 - \langle S \rangle = 1 - \frac{\int_{0}^{r_{max}} \frac{\exp(Pe)\varphi K_c}{\varphi K_c + \exp(Pe) - 1} r^4 f(r) dr}{\int_{0}^{r_{max}} r^4 f(r) dr}$$
(20)

Distribuciones de probabilidad y sus parámetros característicos

La propiedades globales efectivas de transporte dependen del modelo probabilístico supuesto para describir analíticamente la estructura porosa para un mecanismo de transporte aceptado. Los modelos probabilístico para describir la distribución de poros en la estructura porosa a partir de la curva de rechazo son de base empírica. Pueden existir muchos modelos para la predicción de un mismo fenómeno, estos pueden ser empíricos o semi-empíricos y deben ser capaces de reproducir los datos que fueron usados para su construcción (Monahan, 2006).

El problema de la elección de la distribución de probabilidad para un determinado conjunto de datos es de considerable interés en la caracterización de la distribución de poros de las membranas. El modelo probabilístico comúnmente usados para describir la distribución efectiva de poro de las membranas es el Log-normal, (Zydney *et al.* 1994). En este trabajo se consideran las distribuciones log-normal, la gamma, y la Weibull, por ser modelos sesgados a la derecha, fáciles de usar y por tener una forma versátil (Derjani-Bayeh & Rodgers, 2002; Villasana *et al.* 2003).

En la descripción de la distribución efectiva de poros, a pesar de la aparente similitud entre los ajustes de los tres modelos, sigue siendo conveniente evaluar el desempeño relativo de cada uno de ellos, porque las inferencias basadas en el modelo, a menudo involucran la probabilidad de la cola, en la cual el efecto de la hipótesis del modelo es más crítica en la representación del proceso de ultrafiltrado en condiciones de rechazo total.

Los momentos caracterizan las distribuciones probabilísticas y ofrecen medidas que apuntan a diferencias intrínsecas entre ellas. Los mismos proporcionan información sobre las características descriptivas de la distribución. El primer momento es una medida de localización que puede pensarse como el valor central, pero por sí solo no da una adecuada descripción de la forma de dicha distribución. El segundo momento es una medida de la dispersión y es el cuadrado de la desviación estándar de la distribución de los radios. El tercero es una medida de la asimetría relativo a la distribución normal, si la simetría es negativa los datos se extienden más a la izquierda de la media, en cambio si es positiva los datos se extienden más a la derecha. El cuarto es una medida de la concentración de frecuencias alrededor de la media y en la zona central de la distribución en relación a una distribución normal. Este momento, mejor conocido como curtosis, mide cuan pesadas son las colas de la distribución con respecto a la normal que tiene curtosis igual a 3.

Cuando se usan estos tres modelos de probabilidades, los estimados de los cuatro primeros momentos de cada una de las distribuciones de probabilidad se obtienen de las relaciones entre los parámetros optimizados y la definición de momentos. El apéndice A, presenta estos modelos probabilísticos con sus primeros momentos característicos.

Pruebas de bondad de ajuste y validación cruzada

Para evaluar el desempeño relativo de cada uno de los modelos candidatos de la distribución de probabilidad considerados, se adoptaron tres criterios: el análisis de las gráficas de residuos, el error cuadrático medio (*ECM*), la prueba Chi cuadrado (χ^2).

Los residuos son las desviaciones de las observaciones con respecto a la función ajustada que representa el modelo; la gráfica de residuales muestra la diferencia entre los valores calculados y medidos de la variable dependiente en función de los valores medidos. Una distribución aleatoria es indicativo que el modelo es un buen candidato. El error cuadrático medio es un criterio importante y es la sumatoria de la diferencia cuadrática entre los estimadores y el valor real, y se define como:

$$ECM \equiv \sqrt{\frac{\sum (y_{estimado} - y_{real})^2}{N - 1}}$$
(21)

El modelo con menor *ECM* es el modelo de menor varianza y sesgo, y estadísticamente es un buen candidato.

La prueba χ^2 es una prueba que se considera no paramétrica, y representa la diferencia entre la distribución real y la estimada por el modelo (Evans & Rosenthal, 2005). El valor de esta estadística se determina por la ecuación:

$$\chi^2 = \sum_{i} \frac{(estimado_i - real_i)^2}{real_i}$$
(22)

Mientras mayor es el valor de χ^2 , menos probable es la hipótesis de que la estimación es la correcta. En cambio si el valor de χ^2 es cercano a cero, ambas distribuciones están cercanas. (Evans & Rosenthal, 2005).

También se usó la validación cruzada, en la cual un grupo de datos se divide en subconjuntos, con uno de los subconjunto se hace el análisis deseado y con los otros se valida o se confirma el análisis inicial realizado (Kohavi, 1995). De esta manera se intenta comprobar cuán sensible es cada uno de los modelos a la falta de datos. Se puede decir que un modelo es más robusto que otro cuando es menos sensible y más estable a un cambio en las variables, ya sean experimentales o del modelo en sí (Saltelli *et al.* 2005).

PROCEDIMIENTO

La distribución efectiva del tamaño de poro de la estructura porosa de las membranas de ultrafiltración se obtuvo a partir de seis conjuntos de datos de perfil de rechazo previamente publicados (Derjani-Bayeh & Rodgers, 2002), que corresponden a seis diferentes sistemas que se muestran en la tabla 2.

Caso	РМС	Membrana	Soluto	# datos
1	100.000	Polisulfonato	Dextran Mezcla	15

Tabla 2. Datos para perfil de rechazo de distintos casos (Derjani-Bayeh & Rodgers, 2002).

2	100.000	Acetato	Dextran Mezcla	18
3	10.000	Polisulfonato	Dextran Mezcla	12
4	10.000	Polisulfonato	PEG Único	18
5	10.000	Polisulfonato	PEG Mezcla	19
6	30.000	Polisulfonato	PEG Mezcla	15

El conjunto de casos representa un diseño experimental donde se manipula un factor a la vez, como el PMC o peso molecular de corte o capacidad de separación nominal, el tipo de membrana, el tipo de soluto y finalmente la dispersión de soluto (mezcla o único). Esto permite hacer comparaciones con el modelo empírico probabilístico usado para describir la distribución de poros efectiva y su efecto sobre las propiedades globales de transporte.

Los datos experimentales previamente publicados se ajustan al coeficiente promedio teórico de rechazo para la estructura porosa de la membrana, que se determina con la ecuación (20).

Para determinar los parámetros de las funciones de la densidad de probabilidad f(r) se realiza una regresión no lineal usando la subrutina *lsqcurvefit* de Matlab. Este algoritmo ajusta los datos al modelo por medio de rutinas robustas de optimización no lineal (The MathWorks, 2007). Se utiliza la subrutina *nlparci* de Matlab para calcular el intervalo de confianza (basado en la distribución asintótica normal) de dichos parámetros. Este algoritmo retorna el rango de los parámetros, con un 95% de confianza. El proceso de validación cruzada utilizado en este trabajo se realizó quitando un punto a la vez para cada uno de los modelos, funciones y casos. Luego de eliminar el punto, se hace la regresión no lineal y se determinan los parámetros para cada función de distribución.

RESULTADOS

Regresión No Lineal y Bondad de Ajuste

Para todos los casos, los modelos probabilísticos estudiados acoplados al modelo de transporte restringido ajustan estadísticamente bien los datos experimentales de la curva de rechazo. Los resultados son corroborados con las pruebas del error cuadrático medio, *Ecm*, normalizado dentro del 5% y la prueba χ^2 dentro del 10%.

Como ejemplo, se discute el caso 5, la membrana de polisulfonato de 10.000 PMC. En la figura 3, se observa gráficamente los resultados de los ajustes con los parámetros optimizados de cada una de las distribuciones una mezcla de PEG y una membrana de Polisulfonato con PMC de 10.000. En la tabla 3 se reportan los parámetros y sus respectos intervalos de confianza al 95%. La función Log-normal de distribución efectiva de poros permite parámetros negativos, porque los dos primeros momentos son funciones exponenciales, manteniendo el sentido físico de valores positivos del radio efectivo promedio y la desviación estándar.



Figura 3. Ajuste del coeficiente promedio teórico de rechazo (ecuación 20) con los diferentes modelos de densidad de probabilidad a los datos experimentales del caso 5, la curva de rechazo generada con una mezcla de PEG y una membrana de Polisulfonato con PMC de 10.000.

El parámetro de forma, η_g , de la función Gamma de distribución efectiva de poros, es estadísticamente cero, (señalado en color negro oscuro con su intervalo de confianza). Esto implica que la distribución Gamma se reduce a una función de un parámetro que es la exponencial.

Tabla 3. Parámetros de las distribuciones del caso 5, la curva de rechazo generada con una mezcla de PEG y una membrana de Polisulfonato con PMC de 10.000.

F ;(ción Parámetros		IC al 95%	
Funcion			-	+
Gamma	η_{g}	0,1886	-0,0143	0,3914
	$\lambda_{_g}$	0,3352	0,3183	0,3539
Lognormal	μ_{L}	-0,7892	-0,9103	-0,6682
	σ_{L}	0,5064	0,4768	0,5361
Weibull	η_w	0,8558	0,8107	0,9008
	σ_{w}	0,1820	0,1531	0,2108
Exponencial	$\mu_{_{e}}$	0,2776	0,2725	0,2828

Por esta razón se considera la distribución exponencial para describir la distribución efectiva de poros de la membrana a partir del modelo hidrodinámico de transporte, con la densidad parametrizada y sus momentos de acuerdo con la ecuación (23):

$$\begin{cases} f(r) = \frac{1}{\mu_e} e^{-\frac{r}{\mu_e}} \\ \mu_1 = \mu_e \\ \mu_2 = \mu_e^2 \\ \mu_3 = 2 \\ \mu_4 = 6 \end{cases}$$
(23)

Validación Cruzada

,

La validación cruzada indica que todos los modelos probabilísticos escogidos (Log-normal, Weibull, Exponencial) son robustos a la falta de datos. En la tabla 4 se reporta la variabilidad en los parámetros optimizados esta siempre está dentro del rango de los intervalos de confianza de los parámetros optimizados con el conjunto completo de datos.

Existe una tendencia clara a la poca variabilidad del parámetro de la distribución exponencial, menor del 5%. La función Lognormal es la más sensible a la falta de datos o el menos robusto. Su parámetro μ_L , en algunos casos llega a cambiar de signo en los límites de rechazo total y cero rechazo pero siempre dentro del intervalo de confianza, reflejando la importancia de la cola del modelo estadístico.

Tabla 4. Impacto de falta de datos sobre los parámetros de las distribuciones, modelo caso 1, la curva de rechazo generada con una mezcla de Dextran y una membrana de Polisulfonato con PMC de 100.000.

Logn	ormal	Weibull		Exp
$\mu_{\scriptscriptstyle L}$	σ_{L}	η_w	$\sigma_{_w}$	μ_{e}
-0,0460	0,6882	0,4655	0,0532	1,3973
-0,1000	0,6970	0,4622	0,0506	1,4036
0,0649	0,6677	0,5074	0,0928	1,3907
0,0612	0,6685	0,5049	0,0902	1,3934
0,0352	0,6736	0,4947	0,0796	1,3944
-0,1392	0,7145	0,4663	0,0562	1,3944

Análisis de Momentos

La regresión no lineal y las pruebas de bondad de ajuste indicaban que todos los modelos probabilísticos estudiados acoplados al modelo de transporte restringido se ajustan bien a los datos experimentales de la curva de rechazo. En la figura 3, la curva de rechazo es una función de probabilidad acumulada equivalente e indica cualitativamente la calidad de ajuste de los modelos candidatos acoplados al modelo hidrodinámico de transporte. En la figura 4 se muestra la diferencia cualitativa entre los modelos de densidad de probabilidad con parámetros optimizados con regresión no lineal. Cuantitativamente, para poder diferenciar entre los modelos probabilísticos escogidos, se procedió a calcular los cuatro primeros momentos característicos de cada distribución a partir de los parámetros optimizados de la regresión no lineal. De esta manera, se puede verificar si las características intrínsecas de los modelos de distribuciones efectivas de poros son estadísticamente diferentes. Para esto se utilizan las expresiones analíticas de los momentos del apéndice A.

Para considerar si el modelo probabilístico escogido para definir la distribución de poros es una variable tan importante como la capacidad de separación o peso molecular de corte, PMC, se realizó un análisis ANOVA de dos vías, considerando que los momentos dependen tanto de la distribución como de las combinaciones de PMC-soluto-membrana (tabla 5). La importancia de cada factor se midió usando el p-valor calculado a partir de la distribución F, pues de no ser importantes las variables de interés, los momentos estimados serían producto de un error aleatorio que, asintóticamente, aproxima a una distribución normal.



Figura 4. Densidades de Probabilidades, Lognormal, Weibull y Exponencial, con los parámetros optimizados a partir de la curva de rechazo generada del caso 5, con una mezcla de PEG y una membrana de Polisulfonato con PMC de 10.000.

Tabla 5. Análisis ANOVA de dos vías, considerando que los momentos dependen tanto del modelo probabilístico como de las combinaciones de PMC-soluto-tipo de membrana.

Momentos	p-valor distribución	p-valor caso
$\mu_{_1}$	0,0191	0,0136
μ_2	0,1580	0,0130
$\mu_{_3}$	0,0588	0,1905
$\mu_{_4}$	0,1095	0,3028

Es de esperar que el modelo de distribución de poros sea un factor significativo, considerando que PMC es un parámetro muy importante puesto que es una medida de la capacidad nominal de separación de la membrana.

Si la curva sigmoidea presenta una curvatura con alta pendiente, indica que su PMC es alrededor de 10.000 y por lo tanto, que presenta una distribución estrecha de tamaño de poros, indicativo de mejor calidad de separación. Si en cambio la pendiente es poco pronunciada, indicaría un PMC de alrededor de 100.000 y por tanto, una distribución amplia de tamaño de poros, indicativo de una separación difusa.

Los resultados del análisis estadístico ANOVA se presentan en la tabla 5. Indica que los modelos probabilísticos usados para describir la distribución efectiva de poros predicen el primer momento característico con diferencias estadísticamente significativas de un 5 a 10%, (el valor de p es menor a 0,10), que además es equivalente a la significancia estadística del factor de peso molecular de corte. En cuanto al segundo momento, el PMC es más importante que el modelo probabilístico usado para la distribución de poros, mientras que los valores del tercer y cuarto momento indican que los modelos probabilísticos usados para describir la distribución efectiva de poros son de mayor relevancia, mostrado la incidencia de la probabilidad de la cola, en la cual el efecto de la hipótesis del modelo es más crítica en la representación del proceso de ultrafiltrado en condiciones de rechazo total.

Una vez confirmado estadísticamente que el modelo probabilístico para describir la distribución efectiva de poros es tan importante como la separación por peso molecular PMC, se formuló una nueva hipótesis, comparando los casos de membranas que tienen la misma capacidad de separación como lo son los casos 1 y 2 (100.000 PMC) y los casos 3, 4 y 5 (10.000 PMC), usando ahora el análisis ANOVA de una vía. Los resultados se muestran en la tabla 6.

Tabla 6. Análisis ANOVA de una vía, se analiza el efecto de la distribución el estimado de lo momentos para un PMC.

Momentos	100000 PMC casos 1 y 2 p-valor	10000 PMC casos 1, 2 y 3 p-valor
$\mu_{_1}$	0,0945	0,0502
μ_2	0,1328	0,2948
$\mu_{_3}$	0,3421	0,3541
$\mu_{_4}$	0,4229	0,4171

debido a que los valores de p están entre 0,05 y 0,43 para todos los cuatro momentos característicos, se puede afirmar que solo el primer momento depende del modelo probabilístico escogido y, como consecuencia, la predicción del radio promedio efectivo del poro depende del modelo probabilístico.

En la figura 5 se presenta el diagrama de caja (percentiles 25, 50 y 75) de los cuatro momentos característicos. Se puede observar en esta figura que los valores de los cuatro momentos característicos están muy dispersos, de forma más acentuada en la distribución Weibull, lo que hace que el análisis ANOVA muestre los resultados explicados anteriormente, donde el tercer y cuarto momento son estadísticamente iguales para todas las distribuciones. Pero como se ve en los diagramas de caja existen grandes diferencias en la dispersión, lo que trae como consecuencia que el análisis ANOVA no sea concluyente para determinar la igualdad de los momentos dentro de los grupos de casos estudiados.



Figura 5. Diagrama de caja de los cuartos momentos para el caso 1, la curva de rechazo generada con una mezcla de Dextran y una membrana de Polisulfonato con PMC de 100.000.

Influencia del Modelo de la Distribución de Tamaño de Poro en la Predicción del Coeficiente de Filtración

Para determinar el efecto de la distribución en el estimado del coeficiente promedio de filtrado se tomó los parámetros obtenidos de la regresión no lineal con los intervalos de confianza, suponiendo que reflejan el error experimental y se generan las curvas del promedio de filtrado $\langle S \rangle$.

En la figura 6 se muestran las curvas para el caso 1 (membrana de polisulfonato, mezcla de dextrano, con PMC de 100.000) usando los modelos de densidad de probabilidad Lognormal, Weibull y Exponencial, en la que se evidencia que la distribución más sensible al cambio de los parámetros optimizados, es consistentemente la Weibull seguida por la Lognormal. Esto trae como consecuencia que para un radio de soluto dado puedan existir grandes diferencias en la predicción del coeficiente de filtración, *S*, a lo largo de toda la curva, más acentuado en la zona alrededor de la inflexión. La distribución exponencial es la menos sensible a la variación de los parámetros optimizados y por ende la más robusta. La variabilidad del coeficiente de filtración, *S*, para un radio de soluto es muy pequeña, mostrándose dicho comportamiento en todos los casos estudiados.

Para determinar la influencia de las distribuciones sobre la predicción del coeficiente promedio de filtrado, se toman los valores de las curvas más externas de cada distribución en cada caso, como se observa en la figura 7.

Con los datos que generan las curvas de filtrado promedio, alrededor del punto la inflexión de la curva de rechazo, se escoge un rango de radio de soluto entre 5nm a 15nm para los casos 1 y 2 (PMC 100.000), entre 2,5nm y 7,5nm para el caso 3 (PMC 10.000), entre 1,5nm y 3nm para los caso 4 y 5 (PMC 10.000), y entre 1nm y 2,5nm para el caso 6 (PMC 30.000), se realiza el análisis ANOVA de una vía, entre el caso 1, 2 (PMC 100.000), y caso 3, 4 y 5 (PMC 10.000), se obtiene que el valor de *p* es menor a 0,05 indicando que el coeficiente promedio de filtrado es estadísticamente dependiente del modelo probabilístico escogido para describir la distribución efectiva de poro.



Figura 6. Comparación del efecto de los modelos de densidad de probabilidad (Lognormal, Weibull y Exponencial) sobre el coeficiente promedio de filtrado (datos del caso 1, y la curva de rechazo generada con una mezcla de Dextran usando una membrana de Polisulfonato con PMC de 100.000).



Figura 7. Curvas del coeficiente promedio de filtrado con sus límites superior e inferior de las distribuciones Lognormal, Weibull y Gamma (datos del caso 1, la curva de rechazo generada con una mezcla de Dextran usando una membrana de Polisulfonato con PMC de 100.000).

Influencia del modelo de la distribución de tamaño de poro en la determinación del flujo de solvente

Con los parámetros obtenidos para las distribuciones, se realizó el cálculo del flujo promedio de solvente representado por , que de acuerdo a la ecuación (14) quedó expresado como:

$$\langle J_{solvente} \rangle = \frac{\int \frac{\Delta P \pi r^4}{8 \mu L} f(r) dr}{\int \pi r^2 f(r) dr} = \frac{\Delta P}{8 \mu L} \frac{\int r^4 f(r) dr}{\int r^2 f(r) dr}$$
(14)

Para cada caso, los resultados se presentan en la tabla 8.

Tabla 8. Flujos de solventes* con las distintas distribu-ciones, con p-valor de análisis ANOVA.

Caso	Log-normal	Weibull	Exp	p- valor ANOVA
1	16070,5	11837,1	23333,6	0,3099
2	14257,9	12360,1	14594,1	
3	2265,1	1748,3	3721,0	0,6901
4	1478,0	1320,2	1369,1	
5	958,2	824,9	925,0	
6	1688,8	1464,9	1360,6	-

El análisis ANOVA de una vía, entre el caso 1, 2 (PMC 100.000), y caso 3, 4 y 5 (PMC 10.000) indica que no hay diferencias significativas en la predicción del flujo promedio de solvente con diferentes distribuciones, debido a que existen grandes diferencias en la dispersión, lo que trae como consecuencia que el análisis ANOVA no sea concluyente para determinar la diferencias dentro de los grupos de casos estudiados.

CONCLUSIONES

Las pruebas de bondad de ajuste y regresión no lineal son suficientes para considerar que un modelo es bueno, sin embargo, no se pudo emitir un criterio inmediato sobre la validez del modelo. No se pudo discriminar entre los modelos candidatos de probabilidad de densidades. Para todos los modelos probabilísticos acoplados al modelo de transporte restringido se obtuvo un buen ajuste. El error cuadrático medio, *Ecm*, normalizado está dentro del 5% y la prueba χ^2 dentro del 10%.

Con el intervalo de confianza de los parámetros que se obtienen de la regresión no lineal, se logró descartar la distribución gamma debido a que su parámetro de forma es estadísticamente cero.

La prueba de validación cruzada mostró que todas las distribuciones son poco sensibles a la falta de un dato.

Se estudio los efectos del peso molecular de corte y del modelo de densidad de probabilidad en forma simultánea y con una significancia de 0,10, con el análisis de varianza (ANOVA) de dos vías (también llamado análisis de dos factores de variación). Se constató que el modelo probabilístico escogido para definir la distribución efectiva de poros es una variable tan importante como la capacidad de filtración de la membrana o peso molecular de corte (PMC).

Al comparar las diferentes modelos probabilísticos para describir la distribución de poros efectiva en las membranas con un mismo peso molecular de corte (PMC), usando el análisis de la varianza (ANOVA) de una vía, se constató que los modelos probabilístico introducen una diferencia estadística significativa en la predicción del primer momento o radio efectivo promedio de separación y el coeficiente promedio de filtrado.

Se constató que la predicción del flujo promedio de solvente no dependen del modelo probabilístico elegido. Los diagramas de caja mostraron que dichos valores son muy dispersos. Por lo tanto, el análisis ANOVA no contribuyó a diferenciar entre modelos probabilísticos.

REFERENCIAS

DEEN, W. M. (1987). *Hindered Transport of Large Molecules in Liquid-Filled Pores*. AIChE J., 33, pp. 1409-1425.

- DERJANI-BAYEH, S. Y RODGERS, V. (2002). Sieving variations due to the choice in pore size distribution mode, J. Membrane Sci., Vol. 209, pp. 1-17.
- EVANS, M. Y ROSENTHAL, J. (2005). *Probabilidad y estadística: la ciencia de la incertidumbre*, Ciudad de México, Ed. Reverté, pp. 53-145.
- GRANATH, K. A. (1958). Solution properties of branched dextrans. J. Colloid. Sci., 13, pp. 308-328.
- KOHAVI, R. (1995). A Study of Cross-Validation and Bootstrap for accuracy estimation and model Selection, Computer Sci. Departament, Stanford University.
- MICHAELS, A. S. (1980). Analysis and Prediction of Sieving Curves for Ultrafiltration Membranes: A Universal Correlation?. Sep. Sci. Tech., 15, pp. 1305-1322.
- MONAHAN, A. (2006). The probability distribution of sea surface wind speeds. part I: theory and seawinds observations. J. Climate, Vol. 19, pp. 497-520.
- MYERS, R. Y WALPOLE, R. (1999). *Probabilidad y estadística para ingenieros*. México, Pearson Educación, pp. 161-179.
- NIST/SEMATECH, *e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook consultado el 28 de junio de 2008.
- TKACIK, G. AND MICHAELS, S. (1991). A Rejection Profile Test for Ultrafiltration Membranes & Devices. Bio/ Technology., 9, pp. 941-946.
- TAM, C. M. AND TREMBLAY, A. Y. (1991). Membrane Pore Characterization-Comparison Between Single and Multicomponent Solute Probe Techniques. J. Memb. Sci., 57, pp. 271-287.
- SALTELLI, A., RATTO, M., TARANTOLA, S., Y CAMPOLONGO, F. (2005). Sensitivity Analysis for Chemical Models. ACS, Vol. 105, pp. 2811-2827.

THE MATHWORKS. "Matlab Documentation", 1984-2007.

VILLASANA, M., RODGERS, V. G. J. Y DERJANI-BAYEH, S. (2003). Effect of the method of analytically extracting the membrane porous structure in predicting simulated average sieving coefficient in UF membrane Publicado en las actas del 6to. Congreso Interamericano de Computación Aplicada a la Industria de Procesos, CAIP 2003, Puebla México", pp. 149-152.

- WAGNER, J. (2001). Membrane Filtration Handbook Practical Tips and Hints. Cambridge, Osmonics, 2001, pp. 5-38.
- ZYDNEY, A. L., AIMAR, P., MIERELES, M., PIMBLEY, J.M., AND BELFORT, G. (1994). Use of the log-normal probability density function to analyze membrane pore size distributions: functional forms and discrepancies. J. Memb. Sci., 91, pp. 293-298.

Apéndice A

A continuación se describen las funciones de distribuciones de probabilidad usadas para describir la morfología de la membrana (NIST/SEMATECH, 2008; Monahan, 2006).

Lognormal

$$f(r) = \frac{1}{r\sigma L\sqrt{2\pi}} e^{\frac{-(\ln(r)1 - \mu_l)^2}{2\sigma_L^2}}$$

$$\mu_{1} = e^{\mu L + 0.5\sigma L^{2}}$$

$$\mu_{2} = e^{2\mu L + \sigma L^{2}} (e^{\sigma L^{2} + 2\mu L} - 1)$$

$$\mu_{3} = (e^{\sigma L^{2}} + 2)\sqrt{e^{\sigma L^{2}} - 1}$$

$$\mu_{3} = e^{4\sigma L^{2}} + 2e^{3\sigma L^{2}} + 3e^{2\sigma L^{2}} - 6e^{2\sigma L^{2}}$$

Gamma

$$f(r) = \frac{1}{\lambda_g^{\eta_s} \Gamma(\eta_g)} \left(\frac{r}{\lambda_g}\right)^{\eta_s - 1} e^{-\frac{r}{\lambda_g}}$$
$$\mu_1 = \eta_g \lambda_g$$
$$\mu_2 = \eta_g^{-2} \lambda_g$$
$$\mu_3 = \frac{2}{\sqrt{\eta_g}}$$
$$\mu_4 = \frac{6}{\eta_g}$$

Donde la cantidad $\Gamma(.)$ se conoce como función gamma definida por

$$\Gamma(\eta_g) = \int_0^\infty e^{-t} t^{\eta_g - 1} dt$$

Weibull

$$f(r) = \frac{\eta_w}{\sigma_w} \left(\frac{r}{\sigma_w}\right)^{\eta_w - 1} e^{-\left(\frac{r}{\sigma_w}\right)^{\eta_w}}$$
$$\mu_1 = \sigma \Gamma \left(1 + \frac{1}{\eta}\right)$$
$$\mu_2 = \sigma^2 \left(\Gamma \left(1 + \frac{2}{\eta}\right) \Gamma^2 \left(1 + \frac{1}{\eta}\right)\right)$$
$$\mu_3 = \frac{\Gamma \left(1 + \frac{3}{\eta}\right) - 3\Gamma \left(1 + \frac{1}{\eta}\right) \Gamma \left(1 + \frac{2}{\eta}\right) + 2\Gamma^3 \left(1 + \frac{1}{\eta}\right)}{\left(\Gamma \left(1 + \frac{2}{\eta}\right) - \Gamma^2 \left(1 + \frac{1}{\eta}\right)\right)^{\frac{3}{2}}}$$
$$\mu_4 = \frac{\Gamma \left(1 + \frac{4}{\eta}\right) - 4\Gamma \left(1 + \frac{3}{\eta}\right) \Gamma \left(1 + \frac{1}{\eta}\right) + 6\Gamma \left(1 + \frac{2}{\eta}\right) \Gamma^2 \left(1 + \frac{1}{\eta}\right)}{\left(\Gamma \left(1 + \frac{2}{\eta}\right) - \Gamma^2 \left(1 + \frac{1}{\eta}\right)\right)^2} - 3$$