

# APLICACIÓN DE METODOS NUMERICOS PARA EL DESARROLLO DE MODELOS PREDICTIVOS DE PROPIEDADES DE GRANOS

C. A. Sologubik<sup>1,3</sup>, M. R. Ordóñez<sup>2,3</sup>, L. Campañone<sup>4</sup>, M. C. Gely<sup>3✉</sup>, A. M. Pagano<sup>3✉</sup>

<sup>1</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

<sup>2</sup>Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

<sup>3</sup>TECSE, Depto. de Ingeniería Química, Facultad de Ingeniería, UNICEN, Av. del Valle 5737 (7400) Olavarría, Argentina, E-mail: cgely@fio.unicen.edu.ar, apagano@fio.unicen.edu.ar

<sup>4</sup>CIDCA (CONICET La Plata - UNLP), Calle 47 y 116. La Plata, Argentina

**Área Temática:** Aplicaciones de la Matemática

**Palabras Claves:** métodos numéricos, mínimos cuadrados, propiedades físicas.

## RESUMEN

En este trabajo se han aplicado métodos de regresión lineal y no lineal por mínimos cuadrados para estudiar el efecto del contenido de humedad sobre propiedades físicas de granos de Maíz Colorado Duro *Hibrido NT 320* y de Cebada Cervecera variedad *Scarlett*, mediante herramientas computacionales de los entornos MATLAB® y SYSTAT®, evaluándose la correlación entre los valores predichos por diversos modelos matemáticos y los datos experimentales a través de la cuantificación del error. En particular se presentan los resultados obtenidos del análisis basado en métodos numéricos de datos experimentales de la Densidad Real, Densidad Aparente, Esfericidad porcentual, Porosidad porcentual, Ángulo de Reposo, Peso de 1000 granos y Diámetro Equivalente.

## INTRODUCCIÓN

Mínimos cuadrados es una técnica de análisis numérico encuadrada dentro de la optimización matemática en la que, dados un conjunto de pares (o ternas, etc.), se intenta encontrar la función que mejor se aproxime a los datos (un "mejor ajuste"), de acuerdo con el criterio de mínimo error cuadrático (Abdi, 2003). La técnica de mínimos cuadrados se usa comúnmente en el ajuste de curvas. La enseñanza de los Métodos Numéricos actualmente se encuentra íntimamente ligada a la construcción de sistemas matemáticos de alta precisión y exactitud. La mayor parte de los ingenieros depende de herramientas que requieren de aproximaciones cada vez más exactas y precisas. Es por lo tanto deseable que a la par de la enseñanza teórica de los distintos métodos, se aplique un lenguaje de programación de alto nivel que permita escribir directamente todos los algoritmos matemáticos que se van desarrollando. MATLAB® es un entorno de computación y desarrollo de aplicaciones totalmente integrado, orientado para llevar a cabo proyectos en donde se encuentren implicados elevados cálculos matemáticos y la visualización gráfica de los mismos. Este programa integra análisis numérico, cálculo matricial, procesamiento de señal y visualización gráfica en un entorno completo donde los problemas y sus soluciones son expresados del mismo modo en que se escribirían racionalmente, sin necesidad de hacer uso de la programación tradicional (Trombini et al., 2010). SYSTAT® es un poderoso software de estadística que contiene los procesos y aplicaciones necesarios para cualquier procedimiento estadístico que se requiera para un análisis eficiente de datos. Contempla las características de la estadística más elemental hasta la más compleja, utilizando los más sofisticados algoritmos. A diferencia de MATLAB, SYSTAT es un software estadístico de fácil comprensión, manejo y aplicación que permite resolver problemas complejos de todo tipo sin necesidad de ser altamente instruido en programación (SYSTAT, 2010).

Particularmente, en el estudio de procesos tecnológicos aplicados a granos de cereales y oleaginosos (tales como el secado y la aireación) resulta de fundamental importancia el desarrollo de modelos matemáticos que permitan predecir con precisión las propiedades físicas de los materiales. El tamaño de la semilla depende de la variedad (genotipo) y de las condiciones ambientales (humedad y temperatura). Por esta razón el conocimiento de las propiedades físicas y de transporte son indispensables para el adecuado diseño del equipamiento para el manejo, transporte y acondicionamiento de los granos (Kachru et al., 1994). La densidad real del grano es la relación entre una masa conocida de granos y el volumen real ocupado por los mismos. La densidad aparente es la relación entre una masa de granos y el volumen total. Conocer estas propiedades permite determinar el volumen requerido para el almacenamiento, el comportamiento durante el manejo, el transporte y la aireación y el proceso de secado, ya que afectan sobre la velocidad de secado del grano (McCabe et al., 1991; Brooker et al., 1992). La porosidad del lecho de granos se define como la fracción del espacio en un lecho de granos que no es ocupada por granos (Mohsenin, 1970). La necesidad de conocer las propiedades de fricción de materiales granulares como granos y semillas ha tenido un gran reconocimiento en el diseño de los silos y otras estructuras de almacenamiento. El ángulo de reposo es útil tanto en el diseño de equipos para el flujo de sólidos como el diseño de estructuras para el almacenamiento de los mismos (Kachru et al., 1994).

En el presente trabajo se aplica el método de regresión por mínimos cuadrados para el análisis de propiedades físicas de granos de Cebada Cervecera (*Scarlett*) y Maíz Flint utilizando como herramientas computacionales los programas MATLAB y SYSTAT.

## MATERIALES Y MÉTODOS

Se utilizaron granos de Maíz Colorado Duro *Hibrido NT 320* y de Cebada Cervecera variedad *Scarlett*. A fin de estudiar la dependencia de las propiedades físicas del grano con el contenido de humedad, las muestras de

maíz se acondicionaron a diferentes niveles de contenido de humedad, abarcando el rango típico comprendido entre la humedad máxima de cosecha (25% b.h., base húmeda) y la humedad de seguridad (14,5% b.h.). Para el estudio de las propiedades de la cebada cervecera, los granos se acondicionaron abarcando el rango desde la humedad de cosecha (13 %b.s.) hasta el máximo contenido de humedad empleado en el proceso de industrialización de malta (45 % b.s.).

Una breve descripción de las técnicas utilizadas para el estudio de caracterización de los granos es la siguiente:

- *Densidad real* ( $\delta_{real}$ )

Se determinó por picnometría líquida con xileno usando un picnómetro tipo Beckman (Kachru, 1994).

- *Densidad Aparente* ( $\delta_{aparente}$ )

Se determinó usando una balanza estándar de peso hectolítrico (Balanza Schopper) (Kachru, 1994).

- *Peso de 1000 granos* ( $P_{1000}$ )

Se determinó pesando la masa en balanza electrónica (precisión 0,001 g) (Kachru, 1994).

- *Angulo de reposo* ( $AR$ )

Se determinó evaluando el desplazamiento natural por gravedad de los granos utilizando una caja de madera de 300 mm x 130 mm con un panel frontal desmontable (Dulta et al., 1988).

- *Porosidad* ( $\epsilon$ )

Con los valores de densidad real y aparente se calculó la porosidad ( $\epsilon$ ) del lecho de granos, con la siguiente expresión (Gupta y Das, 1997):

$$\epsilon = \left( 1 - \frac{\delta_{aparente}}{\delta_{real}} \right) \quad (1)$$

- *Regresión con mínimos cuadrados*

Para analizar el ajuste de los datos experimentales mediante un modelo lineal se utilizó el método de los mínimos cuadrados, que consiste en minimizar la suma de los cuadrados de las desviaciones ( $L$ ) entre el valor observado ( $y_{medido}$ ) (datos de propiedades físicas) y el valor predicho ( $y_{predicho}$ ) por el método lineal (liberalización del comportamiento de las propiedades físicas):

$$L = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_{medido} - y_{predicho})^2 = \sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2 \quad (2)$$

siendo  $e_i$  el error o residuo y los  $x_i$  el contenido de humedad al que se le efectuó la determinación de las propiedades físicas de los granos. En esta técnica la diferencia de los valores reales y calculados se elevan al cuadrado y se suman, teniendo así las ventajas que las desviaciones positivas y negativas no se cancelan mutuamente.

Entre las muchas ventajas que presenta este criterio, se destaca que se obtiene una línea única para cada cierto conjunto de datos. La técnica que se emplea para calcular los coeficientes  $a_0$  y  $a_1$  que minimizan la Ec. (2) consiste en diferenciar parcialmente  $L$  con respecto a cada uno de los coeficientes y al igualar a cero estas derivadas parciales, con lo cual se obtiene:

$$\frac{\partial L}{\partial \hat{a}_0} = -2 \sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i) = 0 \quad (3)$$

$$\frac{\partial L}{\partial \hat{a}_1} = -2 \sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i) x_i = 0 \quad (4)$$

Luego haciendo  $\sum \hat{a}_0 = n \hat{a}_1$ , siendo  $n$  la cantidad de datos medidos (el comando en MATLAB es “length (n)”) y expresando un sistema de dos ecuaciones normales con dos incógnitas, se obtiene:

$$n \hat{a}_0 + \hat{a}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (4)$$

$$\hat{a}_0 \sum_{i=1}^n x_i + \hat{a}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i \quad (5)$$

Al resolver este sistema de ecuaciones, se encuentra que la estimación de mínimos cuadrados de  $\hat{a}_1$  es:

$$\hat{a}_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{1}{n} \left( \sum_{i=1}^n y_i \right) \left( \sum_{i=1}^n x_i \right)}{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2} \quad (6)$$

Entonces se puede escribir la estimación de mínimos cuadrados de  $\hat{a}_0$  como:

$$\hat{a}_0 = \frac{\sum_{i=1}^n y_i - \hat{a}_1 \sum_{i=1}^n x_i}{n} \quad (7)$$

al resolver la primera de las dos ecuaciones normales para  $\hat{a}_0$ . Esta fórmula para  $\hat{a}_0$  también puede ser expresada como:

$$\hat{a}_0 = \bar{y} - \hat{a}_1 \bar{x} \quad (8)$$

Para simplificar la fórmula para  $\hat{a}_1$  es conveniente introducir la siguiente notación:

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \quad (9)$$

$$S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{1}{n} \left( \sum_{i=1}^n y_i \right)^2 \quad (10)$$

$$S_{xy} = \sum_{i=1}^n y_i (x_i - \bar{x}) = \sum_{i=1}^n y_i x_i - \frac{1}{n} \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right) \quad (11)$$

siendo a  $S_{xx}$  y  $S_{yy}$  la suma corregida de los cuadrados de  $x$  y  $y$ , respectivamente, y  $S_{xy}$  la suma corregida de los productos cruzados de  $x$  y  $y$ . Al emplear esta nueva notación, el estimador de mínimos cuadrados de la pendiente ( $a_1$ ) es:

$$\hat{a}_1 = \frac{S_{xy}}{S_{xx}} \quad (12)$$

### Error estándar en la estimación

El error estándar de la estimación designado por  $SC_E$  mide la diferencia entre las observaciones  $y_i$  y el correspondiente valor predicho  $\hat{y}_i$ , es decir  $e_i = y_i - \hat{y}_i$ , el cual se denomina residuo. La suma de los cuadrados de los residuos, o la suma de cuadrados del error, sería:

$$SC_E = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

Una fórmula de cálculo más conveniente para  $SC_E$  puede encontrarse al sustituir el modelo ajustado  $\hat{y}_i = \hat{a}_0 + \hat{a}_1 x_i$  en la Ec. 12 y simplificar. El resultado es:

$$SC_E = \sum_{i=1}^n y_i^2 - n\bar{y}^2 - \hat{a}_1 S_{xy} \quad (14)$$

y si se deja que  $\sum_{i=1}^n y_i^2 - n\bar{y}^2 = \sum_{i=1}^n (y_i - \bar{y})^2 \equiv S_{yy}$ , entonces es posible escribir  $SC_E$  como:

$$SC_E = S_{yy} - \hat{a}_1 S_{xy} \quad (15)$$

El valor esperado de la suma de cuadrados del error  $SC_E$  es  $E(SC_E) = (n-2)\sigma^2$ . Por tanto,

$$\hat{\sigma}^2 = \frac{SC_E}{n-2} = MC_E \quad (16)$$

es un estimador insesgado de  $\sigma^2$ .

### Coefficiente de determinación

El coeficiente de determinación indica el porcentaje del ajuste que se ha conseguido con el modelo lineal, es decir el porcentaje de la variación de la propiedad física que se explica a través del modelo lineal que se ha estimado, o sea a través del comportamiento de  $MC$  (contenido de humedad %). A mayor porcentaje del coeficiente de determinación mejor resulta el modelo para predecir el comportamiento de las distintas variables de las propiedades físicas. El coeficiente de determinación se define como:

$$r^2 = \frac{SC_R}{S_{yy}} = \frac{a_1 S_{xy}}{S_{yy}} \quad (17)$$

donde  $SC_R$  es la suma de regresión de cuadrados y  $S_{yy}$  es la suma total de cuadrados corregida para  $y$ .

### Coefficiente de correlación lineal

El coeficiente de correlación lineal  $r$  mide de qué manera la recta de los mínimos cuadrados se ajusta a los datos muestrales. Si la variación total se explica totalmente por la recta de regresión, es decir si  $r^2 = 1$  ó  $r = \pm 1$  existe una correlación lineal perfecta. Si  $r$  es positivo y tiende a aumentar con  $x$ , la pendiente de la recta de los mínimos cuadrados es positiva, en tanto si  $r$  es negativo y tiende a disminuir con  $x$ , la pendiente de la recta es negativa. El coeficiente de correlación lineal se define como:

$$r = \frac{\sum_{i=1}^n y_i (x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \sqrt{r^2} \quad (18)$$

En base al objetivo del trabajo, la aplicación de los métodos numéricos, empleando como herramientas computacionales los programas MATLAB y SYSTAT, se creó en el primero un archivo ejecutable "lin\_reg.m", con el propósito de obtener una matriz de gráficos para cada uno de los comportamientos analizados, donde se evidencian los datos experimentales y su ajuste a la linealización, además de cada uno de los coeficientes para la cuantificación del error en la regresión lineal.

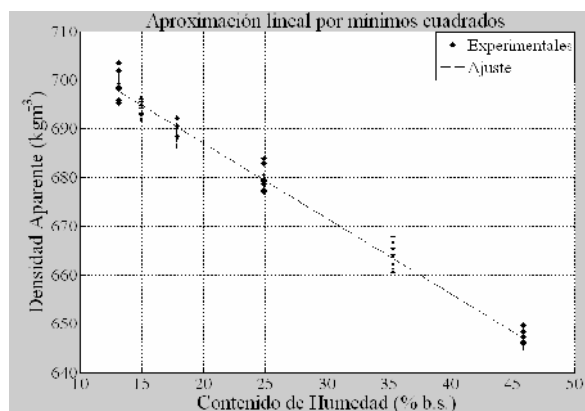
En tanto el paquete estadístico SYSTAT se empleó para corroborar los datos obtenidos por medio del archivo creado en MATLAB y también para resolver el estudio de regresión de las propiedades que no demostraron un comportamiento lineal.

A continuación se describe el programa completo “lin\_reg.m” de MATLAB:

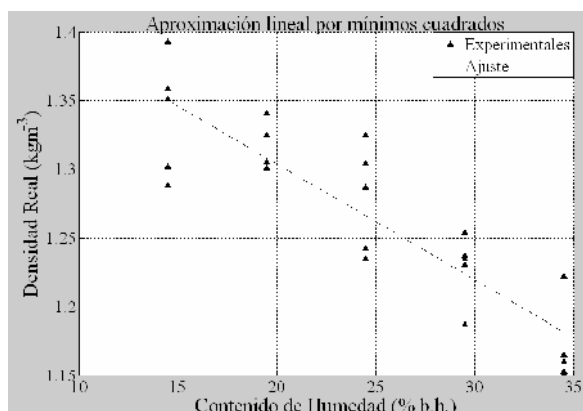
```
function [a,b,r2,r,k]=lin_reg(x,y)
% Número conocido de puntos
n = length(x);
% Iniciación
A=0;
B=0;
C=0;
D=0;
E=0;
% Suma de promedios acumulados
A=sum(x);
B=sum(y);
C=sum(x.*y);
D=sum(x.^2);
E=sum(y.^2);
% Cálculo de los coeficientes de la recta
b=(n*C-A*B)/(n*D-A^2);
a=(B-b*A)/n;
% Cálculo del análisis de regresión
A=b*(C-A*B/n);
E=E-B^2/n;
B=E-A
% Coeficiente de determinación
r2=A/E;
% Coeficiente de correlación
r=sqrt(r2);
% Error estimado std.
k=sqrt(B/(n-2));
% Gráfico
hold on;
plot (x,y,'*','MarkerEdgeColor','r','LineWidth',1);
end
x=x;
y=b*x+a;
plot(x,y,'b');
title('Aproximación lineal por mínimos cuadrados');
```

## RESULTADOS Y DISCUSIÓN

Al ingresar los datos obtenidos de las experiencias de planta piloto, el programa “lin\_reg.m” genera las gráficas (Fig. 1 y 2) para cada una de las propiedades físicas analizadas, mediante la representación de los puntos experimentales junto con las rectas de regresión obtenidas con mínimos cuadrados.



**Fig. 1: Ajuste lineal de la Densidad Aparente de cebada cervecera *Scarlett*.**



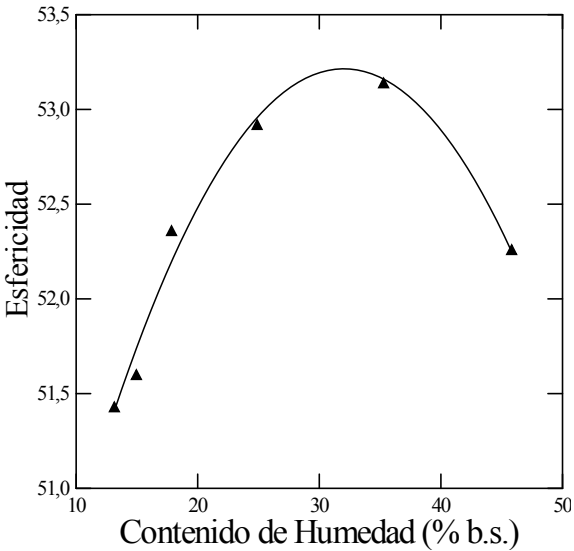
**Fig. 2: Ajuste lineal de la Densidad Real de maíz colorado duro *Hibrido NT 320*.**

Dicho programa devuelve un listado con las constantes de la ecuación de ajuste lineal ( $a$  y  $b$ ), los coeficientes de correlación lineal ( $r$ ), los coeficientes de determinación ( $r^2$ ), los coeficientes estándar en la estimación ( $S_{y/x}$ ) y la suma de los cuadrados de  $y$  ( $S_{yy}$ ) (Tabla 1).

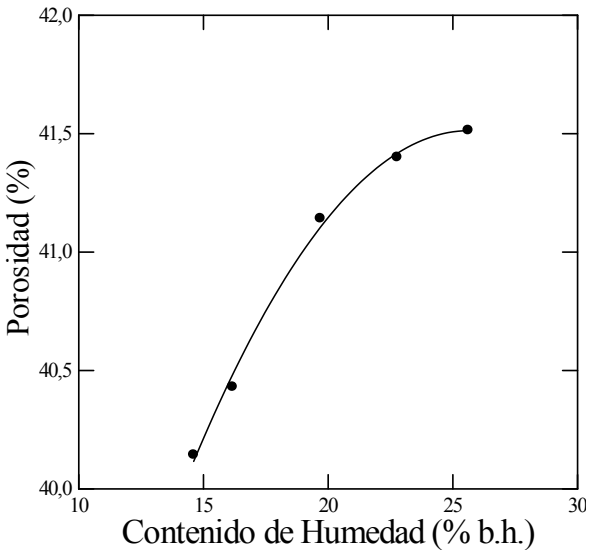
**Tabla 1. Constantes de ajuste lineal y coeficientes para la cuantificación del error.**

Muestra	$a$	$b$	$r$	$r^2$	$S_{y/x}$	$S_{yy}$
Cebada	718,0265	-1,5473	0,9927	0,9855	2,2436	291,9454
Maíz	1,4732	-0,0085	0,8908	0,7934	0,0318	0,0232

En el caso de las propiedades físicas analizadas que presentaron un comportamiento no lineal, se ingresaron los datos obtenidos en el Software SYSTAT realizando un ajuste no lineal. Mediante la función “*Smooth and Plot*” el programa generó las gráficas de los ajustes no lineales (Figs. 3 y 4), representando los puntos experimentales junto con las curvas de regresión obtenidas con mínimos cuadrados.



**Fig. 3: Ajuste curvilíneo de la Esfericidad porcentual ( $\Phi$ ) de cebada cervecera *Scarlett*.**



**Fig. 4: Ajuste curvilíneo de la Porosidad ( $\epsilon$ ) de maíz colorado duro *Híbrido NT 320*.**

Las relaciones existentes entre el contenido de humedad ( $Mc$ ) y las propiedades físicas (Esfericidad y Porosidad) que se muestran en las Figs. 3 y 4, pueden ser representadas por las siguientes ecuaciones polinómicas:

$$\Phi = -0,0051Mc^2 + 0,3244Mc + 48,023$$

considerando  $Mc$  en % b.s.

(19)

$$\epsilon = 0,0052Mc^2 + 0,3486Mc + 35,687$$

considerando  $Mc$  en % b.h.

(20)

con un valor de  $r^2$  de 0,9788 y 0,9967, respectivamente, siendo  $Mc$  el contenido de humedad del grano,  $\Phi$  la Esfericidad porcentual y  $\epsilon$  la Porosidad porcentual.

Por otra parte, se realizó un análisis de varianza (ANOVA) para evaluar la influencia del contenido de humedad sobre las diferentes propiedades físicas estudiadas, considerando un nivel de confianza del 95%. Para ambos granos se observó que el incremento del contenido de humedad de las semillas disminuye la Densidad Real y la Densidad Aparente; en cambio, el Ángulo de Reposo, el Peso de 1000 granos, la Porosidad y el Diámetro Equivalente aumentan con el contenido de humedad de las semillas.

**CONCLUSIONES**

A partir de los resultados alcanzados se puede elaborar como conclusión general la validez del uso de los métodos numéricos y de la herramienta computacional empleada en la resolución de problemas concretos en ingeniería. En particular para los casos planteado se concluye que el modelo de regresión lineal formulado en el Software Matlab permite obtener la recta que mejor se ajusta a la relación del contenido de humedad con las propiedad físicas (Esfericidad y Porosidad), por presentar en ambos casos  $r^2$  superiores a 0,75 (Muñoz Ruiz, 2004). En tanto, para casos más complejos que demostraron un comportamiento No Lineal el empleo del SYSTAT 12 resulto ser efectivo para el ajuste debido a que las ecuaciones polinómicas estimadas presentaron

coeficientes de determinación en el rango de 0,97 a 0,99. Demostrando que dicho programa es un software estadístico de fácil comprensión, manejo y aplicación, que permite resolver problemas complejos sin necesidad de ser altamente instruido en programación.

## REFERENCIAS BIBLIOGRAFICAS

- Abdi H. (2003). Least-squares. *Encyclopedia for research methods for the social sciences*. M. Lewis-Beck, A. Bryman, T. Futing (Eds.), Thousand Oaks (CA): Sage.
- Brooker D B; Bakker-Arkema F W; Hall C W (1992). *Drying and Storage of Grains and Oilseeds*, ISBN 0-442-20515-5, 633.1046 B872, 450pp. Van Nostrand Reinhold, New York.
- Dulta S.K., Nema V.K. and Bharddwaj R.K. (1988). Physical Properties of Gram. *Journal of Agricultural Engineering Research*, 391, 259-268.
- Freud W.E., Miller I. and Miller M. (2000). *Estadística Matemática con Aplicaciones*. Pearson Educación, México, Sexta Edición.
- Gupta R. K and Das S.K.(1997). *Physical Properties of sunflower Seed*. Journal of Agricultural Engineering Research 66:1-8. -McCabe W.L., Smith J.C. and Harriot P. (1991). *Operaciones Básicas de Ingeniería Química*. Ed. Mc Graw-Hill. 4<sup>ta</sup> Edición. ISBN 84-7615-700-2. New York.
- Hines W.W, Montgomery D.C., Goldsman D.M. and Berrer C.M. (2006). *Probabilidad y Estadística para Ingeniería*. Cuarta edición. Compañía Editorial Continental.
- Kachru R. R., Gupta R. K., Alam A. (1994). *Physico-Chemical Constituents & Engineering – Properties of Food Crops*. Editorial Scientific Publishers, Jodhpur.
- Mohsenin N N. *Physical Properties of Plant and Animal Materials*. Gordon and Breach Science Publishers, New York. 1970.
- Muñoz Ruiz D. (2004). “Manual de Estadística”, España, [www.eumed.net](http://www.eumed.net).
- Richard S.R., Mario G.V., Carlos M. (2007). *Hydration of grain kernels and its effect on drying*. Área de Ingeniería Química, Universidad Autónoma Metropolitana-Iztapalapa, D.F., México.
- Trombini G., Córdoba C. A., Rosa M. A. (2010). Aplicación de los métodos numéricos utilizando el software MATLAB® en la determinación de un modelo cinético para la degradación aeróbica de efluentes lácteos. Departamento Química, Facultad Regional Villa María UTN, Villa María, Córdoba, Argentina.
- Wilkinson L. 1990. SYSTAT: The System for Statistics. SYSTAT Inc., Evanston IL, USA.