

USO DE LAS HERRAMIENTAS ESTADÍSTICAS DE EXCEL EN EL ANÁLISIS DE FRECUENCIAS DE CAUDALES MÁXIMOS

J. A. Raynal Villaseñor¹

RESUMEN

Se propone el uso de herramientas estadísticas de Excel® (Excel es una marca registrada de Microsoft, Inc.) en el análisis de frecuencias de caudales máximos anuales. Las metodologías que se proponen usar son todas aquellas capacidades estadísticas y gráficas, que Excel® tiene y que pueden ser aplicadas en la enseñanza del análisis de frecuencias de caudales máximos anuales. En particular, se contiene en el artículo un ejemplo completo de la formulación de la solución del ajuste a la distribución de valores extremos tipo I por medio del uso del método de momentos. Esta tecnología ha sido continuamente utilizada en los cursos de hidrología, a lo largo de varios años, en la Escuela de Ingeniería de la Universidad de las Américas Puebla, con resultados muy satisfactorios, por la facilidad que esta plataforma provee y que hace que el alumno siempre tiene a la vista todos los procesos y tiene también el control de todas las operaciones que se hacen para ajustar las muestras de caudales máximos anuales a la distribución de probabilidad antes mencionada. Los análisis que se presentan en este artículo bajo esta plataforma computacional, son la estimación de los parámetros de la distribución de valores extremos tipo I, sus medidas de bondad de ajuste, la evaluación de los valores de diseño y sus límites de confianza, se incluyen gráficas que muestran el ajuste de los datos a la distribución para diversos métodos de estimación de parámetros y valores de diseño. También se presenta una comparación entre la función de densidad de probabilidad teórica con la empírica. Como conclusiones se pueden establecer que la plataforma utilizada es de gran versatilidad, que los estudiantes pueden comprender y utilizar los conceptos en los que está basada la plataforma computacional y que de una manera totalmente interactiva entre dicha plataforma y el estudiante se induce la curiosidad de éste para explorar todos los aspectos de la tecnología propuesta.

ANTECEDENTES

Un tema de interés primordial en la planificación y diseño de obras hidráulicas es la relacionada con el análisis de frecuencias de caudales máximos anuales. Debido a las características que los valores de diseño, dado que están vinculados a un periodo de retorno y con una probabilidad de no excedencia, el uso de modelos matemáticos conocidos como funciones de distribución de probabilidad es una necesidad. Entre las funciones de distribución de probabilidad más ampliamente usadas para el análisis hidrológico, relacionado con el análisis de frecuencias de caudales máximos anuales, (Kite, 1988, Salas y Smith, 1980, Rao y Hamed, 2000 y Raynal- Villasenor, 2010), son:

- a) normal (NOR)
- b) Log -Normal de dos parámetros (LN2)
- c) Log -Normal de tres parámetros (LN3)
- d) Gamma (GAM)
- e) Pearson III (PIII)
- f) Log- Pearson Tipo III (LPIII)
- g) Valores Extremos Tipo I (VEI)
- h) General de Valores Extremos (GVE)

A la luz de las aplicaciones de computadoras personales en la educación y la formación en todos los ámbitos de la ciencia y la ingeniería, fue desarrollado un grupo de hojas de

¹ Decano de la Escuela de Ingeniería de la Universidad de las Américas Puebla. josea.raynal@udlap.mx.

cálculo para hacerse cargo de los procesos de análisis de frecuencia de caudales máximos anuales, especialmente aplicados para la ingeniería hidrología, pero fácilmente aplicables a otros ámbitos relacionados con los análisis de frecuencias datos máximos. El equipo de cómputo y periféricos requeridos son los que vienen con los equipos de cómputo personal comunes.

METODOLOGÍA

El método de los momentos está diseñado a partir del que se propuso para obtener los momentos de inercia en Estática. Karl Pearson en 1894 adaptó el método de los momentos que se usaba ya en Estática, al considerar a la función de densidad de probabilidad como el cuerpo al cual debían calcularse sus momentos de inercia.

La media y la varianza de la población de la distribución de valores extremos tipo I (VEI) son como sigue, Rao y Hamed (2000):

$$\mu = x_0 + 0.5772 \alpha \quad (1)$$

$$\sigma^2 = \frac{\pi^2}{6} \alpha^2 \quad (2)$$

El valor del coeficiente de asimetría poblacional de la distribución EVI es de 1.1396.

Para la distribución EVI, al utilizar el método de los momentos, se deben igualar las expresiones de la media y la varianza para la población a las de la muestra y después resolver el sistema de ecuaciones simultáneas que resulte, NERC (1975):

$$\mu = \bar{x} \quad (3)$$

$$\sigma^2 = \hat{\sigma}^2 = s^2 \quad (4)$$

Las expresiones para la media y varianza de la muestra son, Haan (1977):

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (5)$$

$$s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (6)$$

Entonces, usando las ecuaciones (1) y (2) en el lado izquierdo de las ecuaciones (3) y (4), y las ecuaciones (5) y (6) en el lado derecho de éstas, se tiene el siguiente sistema de ecuaciones, NERC (1975):

$$x_0 + 0.5772\alpha = \frac{1}{N} \sum_{i=1}^N x_i \quad (7)$$

$$\frac{\pi^2}{6} \alpha^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (8)$$

La solución del sistema de ecuaciones formado por las ecuaciones (7) y (8), produce los estimadores del método de los momentos para la distribución VEI, Rao y Hamed (2000):

$$\hat{x}_0 = \bar{x} - 0.45 \hat{\sigma} = \bar{x} - 0.45 s \quad (9)$$

$$\hat{\alpha} = \frac{\sqrt{6}}{\pi} \hat{\sigma} = \frac{\sqrt{6}}{\pi} s \quad (10)$$

donde $\hat{\mu}$ o \bar{x} es la media de la muestra, $\hat{\sigma}$ o s es la desviación estándar y N es el tamaño de muestra.

Los valores de diseño de la distribución VEI se obtienen al invertir la distribución de probabilidad de la distribución EVI, contenida en la ecuación (11), NERC (1975):

$$F(x) = \exp \left[-\exp \left(-\frac{x - x_{01}}{\alpha_1} \right) \right] \quad (11)$$

Al realizar la operación anterior, el resultado es, Kite (1988):

$$x_T = \hat{x}_0 - \hat{\alpha} \text{Ln} \left[-\text{Ln} (F(x)) \right] \quad (12)$$

donde x_T es el valor de diseño para un cierto valor de la función de distribución $F(x)$.

El término Q_T es más usado en el campo de la ingeniería que el término x_T y esto también se aplica a T_r , el valor del periodo de retorno, en lugar de $F(x)$; por lo tanto, en la siguiente expresión se relaciona el valor de diseño Q_T con un periodo de retorno específico T_r , Raynal-Villasenor (2010):

$$Q_T = \hat{x}_0 - \hat{\alpha} \text{Ln} \left(-\text{Ln} \left(1 - \frac{1}{T_r} \right) \right) \quad (13)$$

donde Q_T es el vlaor de diseño correspondiente a un periodo de retorno específico T_r .

El error estándar de ajuste (EEA) para la distribución VEI tiene la forma siguiente, Kite (1988):

$$EEA = \left[\frac{\sum_{i=1}^N (x_i - y_i)^2}{(N - 2)} \right]^{1/2} \quad (14)$$

Por otra parte, el error medio relativo absoluto (EMRA) es definido como, Jain y Singh (1987):

$$EMRA = \frac{100}{N} \sum_{i=1}^N \left| \frac{(x_i - y_i)}{x_i} \right| \quad (15)$$

donde x_i son los valores históricos de la muestra, y_i son los valores producidos por la función de distribución correspondientes a los mismos periodos de retorno que los de la muestra, y N es el tamaño de la muestra.

La forma general del estimador del método de momentos del error estándar del estimador por ese método para una distribución de dos parámetros es, Kite (1988):

$$S_T^2 = \left(\frac{\partial x_T}{\partial m_1} \right)^2 \text{var}(m_1) + \left(\frac{\partial x_T}{\partial m_2} \right)^2 \text{var}(m_2) + 2 \left(\frac{\partial x_T}{\partial m_1} \right) \left(\frac{\partial x_T}{\partial m_2} \right) \text{cov}(m_1, m_2) \quad (16)$$

y la ecuación (16) puede simplificarse en términos del factor de frecuencia, K_T , como, Kite (1988):

$$S_T^2 = \frac{\mu_2}{N} \left[1 + K_T \hat{\gamma} + \frac{K_T^2}{4} (\hat{\kappa} - 1) \right] \quad (17)$$

Por lo tanto, el estimador del error estándar por el método de los momentos del estimado por ese método, para la distribución VEI es, Kite (1988):

$$S_T^2 = \frac{\sigma^2}{N} \left(1 + 1.1396 K_T + 1.10 K_T^2 \right) \quad (18)$$

donde K_T es el factor de frecuencia dado por, Kite(1988):

$$K_T = - \left[0.45 + 0.7797 \text{Ln} \left(- \text{Ln} \left(1 - \frac{1}{T_r} \right) \right) \right] \quad (19)$$

y usando el factor de frecuencia, la ecuación (18) aún puede simplificarse más de la siguiente manera, Kite (1988):

$$S_T^2 = \frac{\alpha^2}{N} \left(1.1678 + 0.1919 y_T + 1.0999 y_T^2 \right) \quad (20)$$

donde y_T es, Kite(1988):

$$y_T = -Ln \left(-Ln \left(1 - \frac{1}{T_r} \right) \right) \quad (21)$$

Finalmente los límites de confianza pueden expresarse como, Kite (1988):

$$Q_l = Q_T \pm u_\alpha S_T \quad (22)$$

donde Q_l es el límite superior o inferior, definido por el signo más o menos, respectivamente, Q_T es el evento diseño con periodo de retorno T_r y u_α es el nivel de confianza, usualmente para un valor del 95% del nivel de confianza tiene el valor es de 1.96.

DISCUSIÓN DE RESULTADOS

Como ejemplo de aplicación de la metodología descrita en la sección anterior, se procesaron los datos de los caudales máximos anuales de la estación hidrométrica Jaina, Sin., en el periodo (1941-1999), usando el método de los momentos para la distribución VEI. En la Figura 1 se muestran los estadísticos calculados con la función estadísticos descriptivos de Excel[®] (Excel es una marca registrada de Microsoft, Inc.).

En la Figura 2, son mostrados los resultados de la fase de estimación de parámetros y valores de diseño y sus límites de confianza de la estación hidrométrica Jaina, Sin. También se muestran en esta figura los valores de los dos criterios de bondad de ajuste mencionados previamente.

En la Figura 3, se muestran las curvas de frecuencias teóricas y empíricas para la estación hidrométrica Jaina, Sin., obtenidas por diferentes métodos de estimación de parámetros para la distribución VEI.

En la Figura 4, se grafican los límites inferior y superior para los eventos de diseño, producidos por el método de momentos para la distribución VEI en la estación hidrométrica Jaina, Sin.

En la Figura 5, se muestra una comparación entre el histograma de los datos de caudales máximos anuales y la función de densidad de probabilidad para la estación hidrométrica Jaina, Sin.

En la Figura 6, se muestra el comportamiento de los promedios de las calificaciones obtenidas en el curso de Hidrología en el examen donde se evalúa el tema de análisis de

frecuencias de caudales máximos anuales. Se observa un cambio positivo a partir del año 2007 en que fue implantada la metodología descrita en este trabajo.

	A	B
1	ANÁLISIS DE FRECUENCIAS DE CAUDALES MÁXIMOS ANUALES	
2	Estación: Jaina, Sinaloa, México (1941-1999)	
3		
4	<i>Estadísticos Descriptivos</i>	
5		
6	Media	629.9040
7	Error Estándar de la Muestra	86.6215
8	Mediana	433.5000
9	Moda	#N/A
10	Desviación Estándar	665.3521
11	Varianza Muestral	442693.3614
12	Curtosis	16.1116
13	Sesgo	3.4031
14	Rango	4262.3610
15	Mínimo	68.0390
16	Máximo	4330.4000
17	Suma	37164.3365
18	Conteo	59.0000
19	Mayor(1)	4330.4000
20	Menor(1)	68.0390
21	Nivel de Confianza de la Media(95.0%)	173.3917

Figura 1. Estadísticos descriptivos de la estación hidrométrica Jaina, Sin

	A	B	C	D	E	F	G
1	ANÁLISIS DE FRECUENCIAS DE CAUDALES MÁXIMOS ANUALES						
2	Estación: Jaina, Sinaloa, México (1941-1999)						
3	Valores Extremos Tipo I (VEI)						
4	Método: Momentos						
5							
6	Estimación de Parámetros y Pruebas de Bondad de Ajuste						
7	Parámetro de Ubicación =	330.4956	m ³ /s				
8	Parámetro de Escala =	518.7717	m ³ /s				
9	EEA =	282		EMRA =	23		
10							
11	Valores de Diseño y Límites de Confianza						
12		Factor de	Raíz Cuadrada	Límite Inferior	Valores de	Límite Superior	
13	Periodo de Retorno	CDF	Frecuencia	Error Estándar	de Dos Colas	Diseño	de Dos Colas
14	Tr (años)	F(x)	K _T	S _T	95%	Q _T (m ³ /s)	95%
15	2	0.5	-0.1642	285.5108	-39	521	1080
16	5	0.8	0.7195	480.8158	166	1109	2051
17	10	0.9	1.3046	649.4193	225	1498	2771
18	20	0.95	1.8659	820.3734	263	1871	3479
19	50	0.98	2.5923	1047.7680	301	2355	4408
20	100	0.99	3.1367	1220.5669	325	2717	5109
21	500	0.998	4.3947	1623.8474	371	3554	6737
22	1000	0.999	4.9356	1798.2148	389	3914	7438
23	5000	0.9998	6.1908	2204.1634	429	4749	9069
24	10000	0.9999	6.7313	2379.3387	445	5109	9772

Figura 2. Estimación de parámetros, valores de diseño y sus límites de confianza de la estación hidrométrica Jaina, Sin

**Análisis de Frecuencias de Caudales Máximos Anuales
Estación Hidrométrica: Jaina, México (1941-1999)
Distribución de Valores Extremos Tipo I**

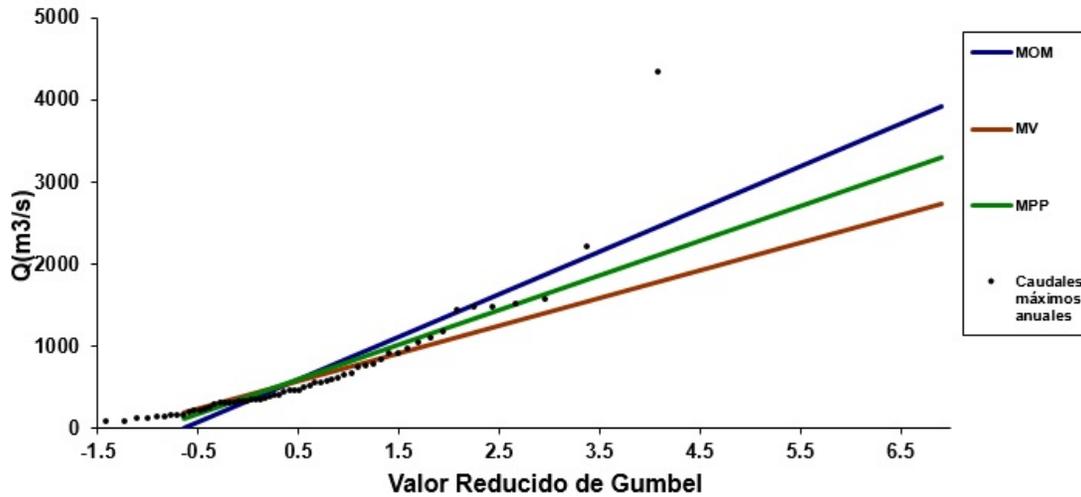


Figura 3. Curvas de frecuencias teóricas y empíricas para la estación hidrométrica Jaina, Sin

**Límites de Confianza
Estación Hidrométrica: Jaina, Sin. (1949-1999)
Distribución de Valores Extremos Tipo I
Método de Momentos**

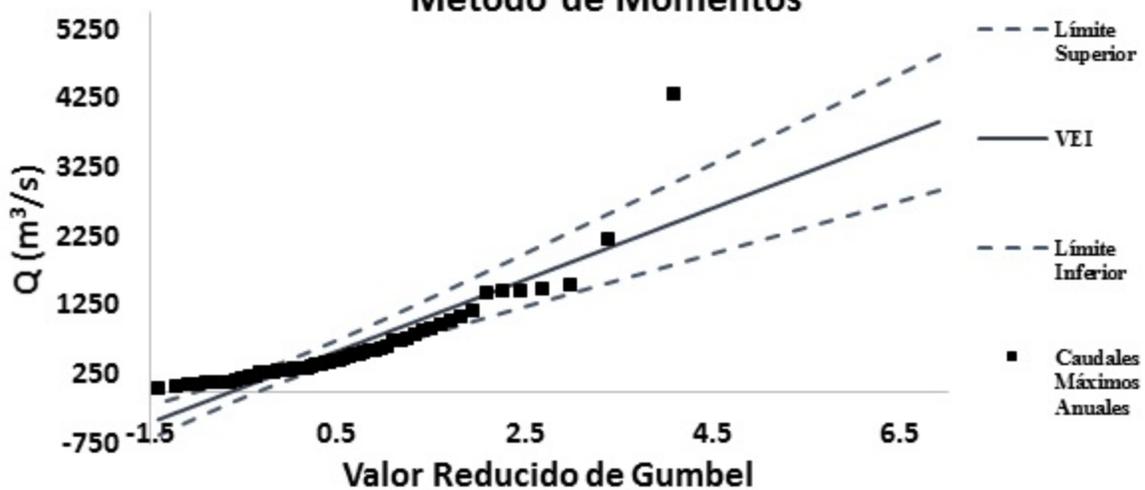


Figura 4. Límites de confianza para la estación hidrométrica Jaina, Sin

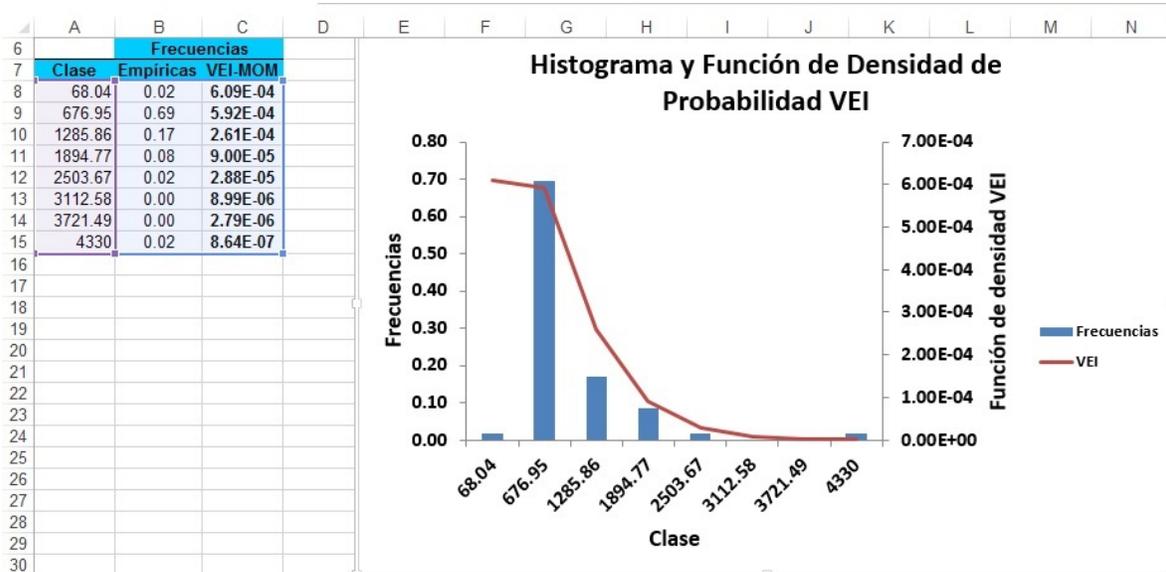


Figura 5. Histograma de los caudales máximos anuales y función de densidad de probabilidad VEI para los datos de la estación hidrométrica Jaina, Sin

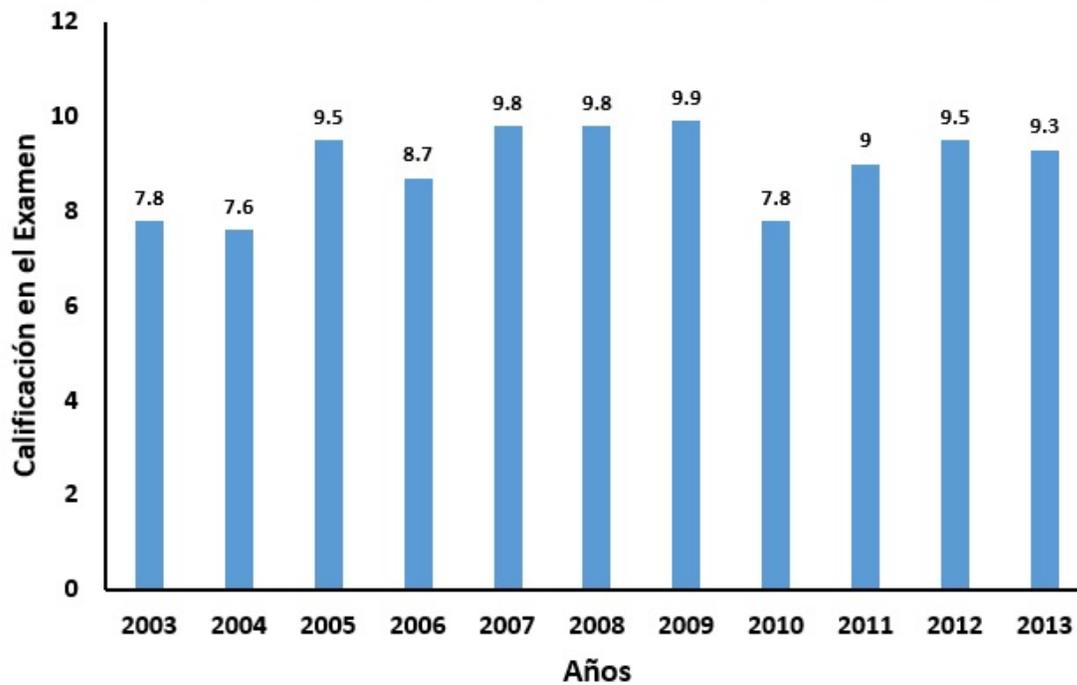


Figura 6. Comportamiento de las calificaciones en el examen que evalúa el tema de análisis de frecuencias de caudales máximos anuales dentro del curso de Hidrología

CONCLUSIONES

Se han presentado algunas de las aplicaciones estadísticas que posee Excel[®], para su aplicación en el análisis de frecuencias de caudales máximos anuales en la educación y el entrenamiento de personas interesadas en esta área del conocimiento hidrológico. Debido a los requisitos mínimos de memoria central y periféricos de asociados a una computadora, ya que procedimiento mostrado hace uso de Excel[®], una herramienta de cálculo estándar, como el principal entorno computacional, esta virtud pone el procedimiento descrito a la disposición de cualquiera que tenga acceso a una computadora personal. Además se han observado resultados positivos en el comportamiento de los promedios de calificaciones obtenidas en los exámenes donde se evalúa el tema de análisis de frecuencias de caudales máximos anuales, que se ha descrito en este artículo.

BIBLIOGRAFÍA

Haan, C. T., 1977, *Statistical Methods in Hydrology*, The Iowa State University Press, Ames, Iowa, E.U.A.

Jain, D. y Singh, V.P., 1987, *Estimating Parameters of EVI Distribution for Flood Frequency Analysis*, Water Resources Bulletin, Vol. 23, No. 1, 59-71

Kite, G.W., 1988, *Flood and Risk Analyses in Hydrology*, Water Resources Publications, Littleton, Colorado, E.U.A.

Natural Environment Research Council (NERC), 1975, *Flood studies report, vol. 1 Hydrological studies*, Whitefriars LTD. Londres. Inglaterra.

Rao, A. R. and Hamed, K. H., 2000, *Flood Frequency Analysis*, CRC Press, Boca Raton, Florida, E.U.A.

Raynal-Villasenor, J. A., 2010, *Frequency Analysis of Hydrologic Extremes*, Lulu.com, E.U.A.

Salas, J. D. and Smith, R., 1980, *Computer Programs of Distribution Functions in Hydrology*, Colorado State University, Fort Collins, Colorado, E.U.A.