



CUADERNOS DE INGENIERÍA

Nueva serie

PUBLICACIONES ACADÉMICAS

9

Año 9 / Número IX / diciembre, 2016.



EDITORIAL
UNIVERSIDAD CATÓLICA DE SALTA
EUCASA

UNIVERSIDAD CATÓLICA DE SALTA
(UCASAL)

CUADERNOS DE INGENIERÍA
N° 9 | 2016

CONSEJO EDITORIAL

DIRECTORA: Dra. Ing. Lía Orosco Segura | Jefa del Departamento de Investigación, Universidad Católica de Salta

Mg. Ing. Néstor Eugenio Lesser | Decano de la Facultad de Ingeniería, Universidad Católica de Salta

Dr. en Ciencias Luis Olsina Santos | Facultad de Ingeniería, Universidad Nacional de la Pampa

Ing. Joan Carles Olmedillas Hernández | Indra Espacio S. A, Barcelona, España

Dr. Ing. Jorge Eugenio Crempien Laborie | Pontificia Universidad Católica de Chile

Dr. Ing. Horacio Colina | École Nationale de Ponts et Chaussées, Francia

Dr. José Germán Viramonte | CONICET, Universidad Nacional de Salta

Cuadernos de Ingeniería es la revista científica de la Facultad de Ingeniería, de la Universidad Católica de Salta, Argentina, creada en el año 2006 y publicada bajo el sello EUCASA (Ediciones Universidad Católica de Salta).

Es una publicación de periodicidad anual, que contiene artículos, ensayos, notas de divulgación y reseñas de libros sobre tópicos directamente relacionados con el campo disciplinar de la ingeniería en todas sus ramas, como así también temas tales como la enseñanza de la ingeniería, el desempeño de la profesión, la ética y los valores en la ingeniería, las interrelaciones entre la ingeniería y la política, la cultura, la economía, la sociedad y el medio ambiente.

«*Cuadernos de Ingeniería*» is the scientific magazine published by «Universidad Católica de Salta», Argentina. It was created in 2006 and is an imprint of EUCASA (Ediciones Universidad Católica de Salta - Universidad Católica de Salta Editions).

It is an annual publication which contains papers, essays and book reviews about topics directly related with the wide field of engineering science, including relationship between engineering and politics, ethic, culture and economy sciences, as well as engineering education and professional matters.

Año 9 | número IX | diciembre 2016

ISSN 2545-7012 (on line)

ISSN 2545-692X (papel)

Domicilio editorial: Campus Universitario Castañares,

4400 Salta - Argentina

Tel./fax: (54-387) 426 8607

e-mail: eucasa@ucasal.edu.ar

web: www.ucasal.edu.ar/eucasa

Imagen de tapa:

Boceto. Leonardo Da Vinci (1452-1519)

CUADERNOS DE INGENIERÍA

Publicaciones Académicas

Nueva Serie

9 | 2016



EDICIONES
UNIVERSIDAD CATÓLICA DE SALTA
SALTA - ARGENTINA

SUMARIO

Editorial	5
Aplicación del Método de Galkerin al análisis de la estabilidad del equilibrio elástico de las columnas	7
Carlos M. Albarracín - Fernando J. Albarracín - Lía E. Orosco Segura	
Análisis de las condiciones locales para introducir transformadores de distribución de núcleos amorfos	17
Soledad Gamarra - Jorge Horacio Giubergia - José Alderete, Manuel Luis Zambrano - Javier Moya	
Estimación de láminas de lluvias máximas diarias y sus parámetros estadísticos en el norte de Argentina	29
Nicolás F. Guillén - M. Cecilia Botelli - Carlos M. García - Carlos Gastón Catalini	
Analysis of Gutenberg-Richter b-value and m_{\max} . Part I: Exact Solution of Kijko-Sellevoll Estimator of m_{\max}	51
Mika Haarala y Lía Orosco	
Analysis of Gutenberg-Richter b-value and m_{\max} . Part II: Estimators for b-value	79
Mika Haarala y Lía Orosco	
Diseño de un taller de introducción a la programación de robots	107
M. Lorena Talamé - Norberto Aramayo - Ana E. Gardel - M. Alicia Pérez Abelleira	
Seguridad de las Redes de Control Industrial – MODBUS/TCP con inspección profunda de paquetes	121
Carlos Said	
Perfil académico- profesional de los autores	129

Editorial

Como cada año, los artículos más relevantes resultantes de la actividad de investigación realizada por los docentes de la Facultad de Ingeniería y por los miembros de los distintos grupos de investigación que integran el Instituto de Estudios Interdisciplinarios de Ingeniería (IEsIIng), son incluidos en los Cuadernos de Ingeniería. Dado que en 2015 se discontinuó la publicación, se inaugura con este número, una nueva serie que abre caminos hacia nuevos desafíos y mayores exigencias de calidad.

Indudablemente, cada artículo es un indicador válido del fortalecimiento científico de cada equipo en particular y de la Facultad en su conjunto.

El equipo que se especializa en **Métodos Numéricos Aplicados a la Ingeniería**, presenta el artículo «Aplicación del Método de Galerkin al análisis de la estabilidad del equilibrio elástico», en el que se ha adoptado como modelo matemático de barras, el de Euler-Bernoulli, y se incluye la variación de la sección y la presencia de un apoyo elástico en la dirección de la longitud del elemento estructural.

Como producto de las investigaciones realizadas por el grupo que estudia el **Riesgo Hidrológico**, el artículo «Estimación de láminas de lluvias máximas diarias y sus parámetros estadísticos en el norte de Argentina», da respuestas a algunos aspectos no estudiados en profundidad. Actualmente, una gran extensión del territorio argentino no cuenta con herramientas normalizadas para predecir los parámetros de diseño hidrológico de las medidas estructurales y no estructurales requeridas para mitigar la amenaza hídrica; por lo que cada proyectista utiliza los datos disponibles y metodologías de predicción diferentes, según su grado de experiencia y criterio.

El equipo que trabaja sobre **Riesgo Sísmico** aborda el problema de la estimación de los parámetros b y M_{\max} con dos artículos: «Analysis of Gutenberg-Richter b -value and m_{\max} . Part I: Exact Solution of Kijko-Sellevoll Estimator of m_{\max} » y «Analysis of Gutenberg-Richter b -value Part II: Estimators for b -value», de gran solidez matemática, proponiendo soluciones exactas y cerradas a su tratamiento, esencial para estudios de peligrosidad sísmica.

Los Cuadernos de Ingeniería también incluyen propuestas de innovación, con artículos que permiten visualizar el futuro con los nuevos materiales y el estudio de sus nano propiedades. En esta línea el **Grupo Interdisciplinario de Materiales** en «Análisis de las condiciones locales para introducir transformadores de distribución de núcleos amorfos», realiza una breve descripción de los AMDT, exponiendo sus ventajas. Se analizan las principales razones que limitan su aplicación, con el objetivo de proponer algunas acciones tendientes a la incorporación de esta tecnología, relacionadas con la difusión del conocimiento y con un marco legal que incentive la producción de los AMDT en empresas locales, que se incluya dentro de las políticas de ahorro

energético vigentes.

El grupo **Robótica** propone el «Diseño de un taller de introducción a la programación de robots», el interés por la robótica entre alumnos de la carrera de Ingeniería en Informática y la ausencia de temas directamente relacionados con la misma en el currículum, junto con la disponibilidad cada vez mayor de robots a costos accesibles, ha motivado el diseño de un taller de introducción a la programación de robots. Esta iniciativa dio origen a valiosos proyectos de extensión.

El artículo «Seguridad de las Redes de Control Industrial – MODBUS/TCP con inspección profunda de paquetes» incluye un tema que puede generar interés para la conformación de nuevos grupos de investigación. Su objetivo es tratar la aplicación y la gestión de las prácticas de seguridad cibernética asociadas con la tecnología de la información y la tecnología operativa en relación con los entornos de redes de producción industrial (SCADA-ICS).

La evolución de la investigación se mide a través de las producciones; esta variedad de temas nos permite inferir que el camino está iniciado y se vislumbra un futuro promisorio. La Facultad debe mantener esta política de fomento a la participación de los docentes en actividades de investigación. Esta es nuestra apuesta.

Mg. Ing. Néstor Lesser
Decano de la Facultad de Ingeniería
Universidad Católica de Salta

Aplicación del Método de Galerkin al análisis de la estabilidad del equilibrio elástico de las columnas

Carlos M. Albarracín¹ – Fernando J. Albarracín¹ – Lía E. Orosco Segura¹

Resumen

La determinación de las cargas críticas que pueden ocasionar inestabilidad en elementos estructurales tipo columnas, es de vital importancia en el diseño estructural, ya que una falla de este tipo puede originar el colapso de una estructura. En esta investigación se ha adoptado como modelo matemático de barras, el de Euler-Bernoulli, se ha incluido la variación de la sección y la presencia de un apoyo elástico en la dirección de la longitud del elemento estructural. En primer lugar se procedió a derivar la Formulación Débil del problema y se aplicó el método de Galerkin para determinar las cargas críticas. A partir de la adopción de espacios de polinomios de dimensión finita, se pudo abordar el estudio de diferentes problemas con condiciones clásicas de apoyo.

Palabras clave: Galerkin - pilotes - suelo

Abstract

The estimation of the critical load that can produce buckling in columns-type elements, is of primarily importance, since it can origin structural collapse. In this work, the beam model of Euler-Bernoulli has been adopted to analyze the problem, and variable cross section and elastic support in the direction of the element, are included.

At first it is derived a weak formulation problem, and afterwards, the Galerkin Method is applied to calculate the critical loads. The adoption of spaces of finite dimension polynomials made possible to handle the solution of different load cases as well as different support conditions.

Keywords: Galerkin - piles - soil

1. Introducción

La determinación de las cargas críticas de pandeo en elementos estructurales tipo columnas, como es sabido, resulta de vital importancia en el diseño estructural. El modelo matemático

¹ Es.I.Ing. Instituto de Estudios Interdisciplinarios de Ingeniería, Universidad Católica de Salta, Argentina.

más usual es aquel que presenta barras con deformación por flexión de acuerdo con el modelo de barra de Euler-Bernoulli. En los problemas clásicos, las columnas pueden presentar diferentes condiciones de apoyo en sus extremos y existe una amplia bibliografía en la cual se detallan los modelos matemáticos y las soluciones exactas correspondientes. En Ingeniería Civil por ejemplo, la posibilidad de falla por pandeo en fundaciones profundas (con pilotes) puede en ocasiones estar subestimada, por lo que es preciso tener en cuenta en el modelo matemático la estructura de fundación y su interacción con el suelo, el que puede ser modelado por medio de resortes. En el caso de diseño de elementos mecánicos esbeltos, la variación de espesor es una práctica muy común, ya que de esa manera se logran componentes mecánicos de poco peso, importante si estos elementos trabajan con movimiento.

El objetivo de este trabajo, es determinar valores de cargas críticas de pandeo y sus correspondientes modos de falla en el caso de barras con espesor variable, restricción elástica en el sentido longitudinal y considerando varias de las condiciones de apoyo clásicas. La obtención de la formulación débil del problema permite la aplicación del método de Galerkin, y con él se pueden determinar valores de cargas críticas que ocasionan inestabilidad del equilibrio elástico.

2. Formulación

Considérese una barra tal como se muestra en la Figura 1, sometida a la acción de una carga P que siempre mantiene una dirección paralela al eje \bar{x} . Esta barra está constituida por un material isótropo y homogéneo, de momento de inercia $I(\bar{x})$ cuya expresión es función de la variación de la sección. La incorporación de una restricción elástica $K(\bar{x})$ lo largo de la barra limita los desplazamientos transversales. Estos desplazamientos se suponen posibles sólo en el plano $\bar{x} - \bar{y}$.

A los efectos ilustrativos, se considera un problema con un empotramiento en la base y libre en el otro extremo.

El problema de contorno en este caso es

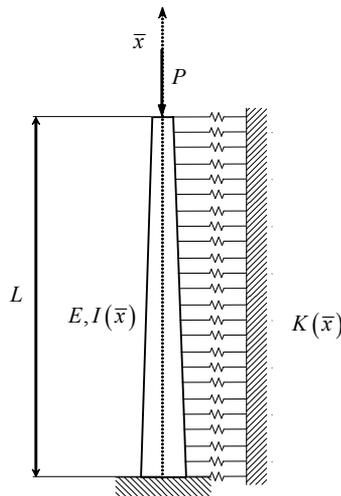


Figura 1.

$$\begin{aligned} \frac{d^2}{d\bar{x}^2} \left[EI(\bar{x}) \frac{d^2 w(\bar{x})}{d\bar{x}^2} \right] + P \frac{d^2 w(\bar{x})}{d\bar{x}^2} + K(\bar{x}) w(\bar{x}) &= 0, \quad \forall \bar{x} \in (0, L), \\ w(\bar{x}) \Big|_{\bar{x}=0} &= 0, \quad \frac{dw(\bar{x})}{d\bar{x}} \Big|_{\bar{x}=0} = 0, \\ \frac{d^2 w(\bar{x})}{d\bar{x}^2} \Big|_{\bar{x}=L} &= 0, \quad \frac{d}{d\bar{x}} \left[EI(\bar{x}) \frac{d^2 w(\bar{x})}{d\bar{x}^2} \right] + P \frac{dw(\bar{x})}{d\bar{x}} \Big|_{\bar{x}=L} = 0. \end{aligned} \quad (1)$$

Al adoptar un cambio de variable de la forma $x = \frac{\bar{x}}{L}$, se tiene el siguiente problema de contorno en forma adimensional

$$\begin{aligned} \frac{d^2}{dx^2} \left[\frac{I(x)}{I_0} \frac{d^2 w(x)}{dx^2} \right] + \alpha \frac{d^2 w(x)}{dx^2} + \beta k(x) w(x) &= 0, \quad \forall x \in \Omega = (0, 1), \\ w(x) \Big|_{x=0} &= 0, \quad \frac{dw(x)}{dx} \Big|_{x=0} = 0, \\ \frac{d^2 w(x)}{dx^2} \Big|_{x=1} &= 0, \quad \frac{I(x)}{I_0} \frac{d^3 w(x)}{dx^3} + \alpha \frac{dw(x)}{dx} \Big|_{x=1} = 0, \end{aligned} \quad (2)$$

en donde $I_0 = I(x) \Big|_{x=0}$, $\alpha = \frac{PL^2}{EI_0}$ y $\beta = \frac{L^4}{EI_0}$.

El espacio apropiado para ubicar el problema de contorno asumiendo

$$k, I \in L^2(\Omega), \text{ es}$$

$$V(\Omega) = \left\{ v \in H^2(\Omega), v \Big|_{x=0} = v' \Big|_{x=0} = 0 \right\}. \quad (3)$$

Para derivar la formulación débil, se aplica el procedimiento habitual, tal como se expone en Atkinson y Han (2001), es decir

$$\int_0^1 \left\{ \frac{d^2}{dx^2} \left[\frac{I(x)}{I_0} \frac{d^2 w(x)}{dx^2} \right] + \alpha \frac{d^2 w(x)}{dx^2} + \beta k(x) w(x) \right\} v(x) dx = 0, \quad (4)$$

para toda $v \in V(\Omega)$. Cuando en (4) se aplica la fórmula de integración por partes dos veces al primer término de la integral, una vez al segundo término y se consideran las condiciones de contorno del problema, se arriba a la siguiente expresión

$$\frac{1}{I_0} \int_0^1 I(x) w''(x) v''(x) dx - \alpha \int_0^1 w'(x) v'(x) dx + \beta \int_0^1 k(x) w(x) v(x) dx = 0. \quad (5)$$

En consecuencia, la formulación débil resulta

$$\begin{cases} \text{Hallar } w \in V(\Omega) \text{ tal que} \\ a(w, v) - \alpha (w', v')_{L^2} = 0, \quad \forall v \in V(\Omega), \end{cases} \quad (6)$$

en donde

$$a(w, v) = \frac{1}{I_0} \int_0^1 I(x) w''(x) v''(x) dx + \beta \int_0^1 k(x) w(x) v(x) dx \quad (7)$$

y

$$(w', v')_{L^2} = \int_0^1 w'(x) v'(x) dx. \quad (8)$$

La formulación es un problema de autovalores y la existencia y unicidad de su solución no son tratadas en este trabajo, pero pueden ser consultadas en el trabajo de Rektorys (1980).

3. Aplicación del método de Galerkin

A partir de la formulación variacional obtenida en la sección anterior, es posible aplicar el método de Galerkin para obtener valores aproximados de los coeficientes adimensionales α de cargas críticas.

Si se considera el espacio de dimensión finita $V^h(0,1) \subset H^2(0,1)$, el planteo de la formulación débil en este espacio resulta

$$\begin{cases} \text{Hallar } w_h \in V^h(\Omega) \text{ tal que} \\ a(w_h, v_h) - \alpha (w_h', v_h')_{L^2} = 0, \quad \forall v_h \in V^h(\Omega). \end{cases} \quad (9)$$

Toda función perteneciente a V^h puede ser expresada como combinación lineal de una base, por lo tanto, w_h, v_h resultan

$$w_h = \sum_{j=1}^N c_j \varphi_j, \quad v_h = \sum_{i=1}^N d_i \varphi_i, \quad (10)$$

en donde las $\varphi_i(x)$ son funciones de la base de V^h y N la cantidad adoptada.

Al reemplazar las expresiones en la y operar algebraicamente, se obtiene el siguiente problema de autovalores

$$(\mathbf{K} - \alpha \mathbf{M}) \mathbf{C} = \mathbf{0}, \quad (11)$$

en donde las matrices \mathbf{K} y \mathbf{M} son simétricas y \mathbf{C} es el vector de los parámetros desconocidos c_j . Los componentes \mathbf{K} y se obtienen de la siguiente forma

$$\mathbf{K}_{i,j} = a(\varphi_i, \varphi_j), \quad (12)$$

$$\mathbf{M}_{i,j} = (\varphi'_i, \varphi'_j)_{L^2}. \quad (13)$$

4. Resultados numéricos

Como ejemplo de prueba, se presenta el análisis de una columna de espesor constante y empotrada en la base y libre en el extremo superior. La base propuesta para el espacio está dada por el conjunto de funciones Los resultados que se muestran en la tabla 1 coinciden con los publicados por Wang et al. (2005).

Tabla 1: Coeficientes α correspondientes a los 4 primeros modos y para diferentes valores de N .

N	α_1	α_2	α_3	α_4
1	3.000000	-----	-----	-----
2	2.485962	32.180705	-----	-----
3	2.467738	23.391255	109.141007	-----
4	2.467404	22.321761	69.404350	265.806484
5	2.467401	22.213852	63.027671	148.205039
6	2.467401	22.206915	61.862571	127.206852
7	2.467401	22.206619	61.701715	122.167158
8	2.467401	22.206610	61.686176	121.097190

Para el siguiente problema, la formulación débil resulta idéntica a la dada en , siempre y cuando las condiciones de contorno sean homogéneas. Para obtener la solución correcta, sólo se debe seleccionar apropiadamente una base para el espacio en donde se define la solución del problema. En este caso se trata de una columna con base empotrada, extremo superior simplemente apoyado, restricción elástica y momento de inercia variable de acuerdo con la expresión en donde y son constantes y es una dimensión de la sección transversal. El valor del parámetro permite modelar las secciones más usuales en ingeniería. A continuación se presentan valores del parámetro de los cuatro primeros modos, los cuales se obtienen adoptando También se muestran las formas modales correspondientes a

Tabla 2: Coeficientes α para $h(x) = 1 + x, \forall x \in [0, 1]$ y $D = 1$.

n	α_1	α_2	α_3	α_4
1	29.448963	86.988418	173.278815	288.328976
2.4	48.301874	142.563149	283.924481	472.400228
3	58.957689	174.037185	346.618134	576.718517
4	80.762914	238.718064	475.599477	791.431245

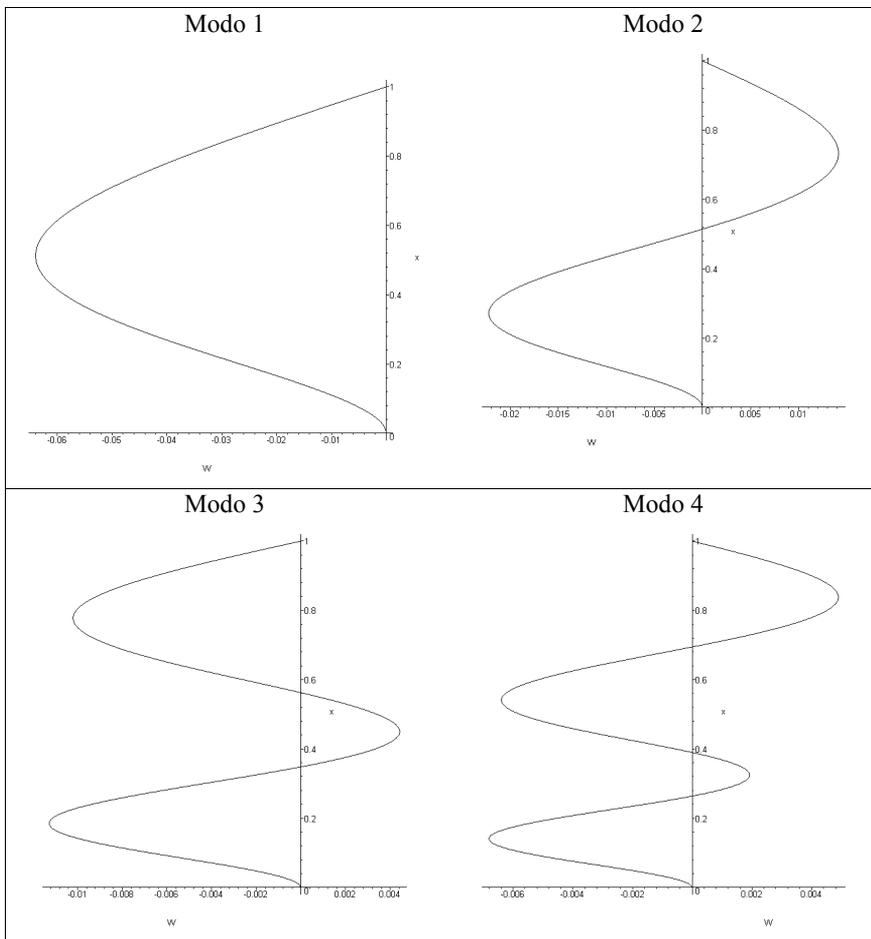


Figura 3. Modos normales de pandeo para $n = 3$.

En el problema que se propone a continuación, una columna simplemente apoyada en ambos extremos, presenta una restricción elástica constante a lo largo de toda su longitud, con momento de inercia y en la figura 4 se muestra el primer modo y en ella se puede observar que con el incremento de $k(x)$ se producen consecuentes aumentos en el número de las semi-ondas del primer modo.

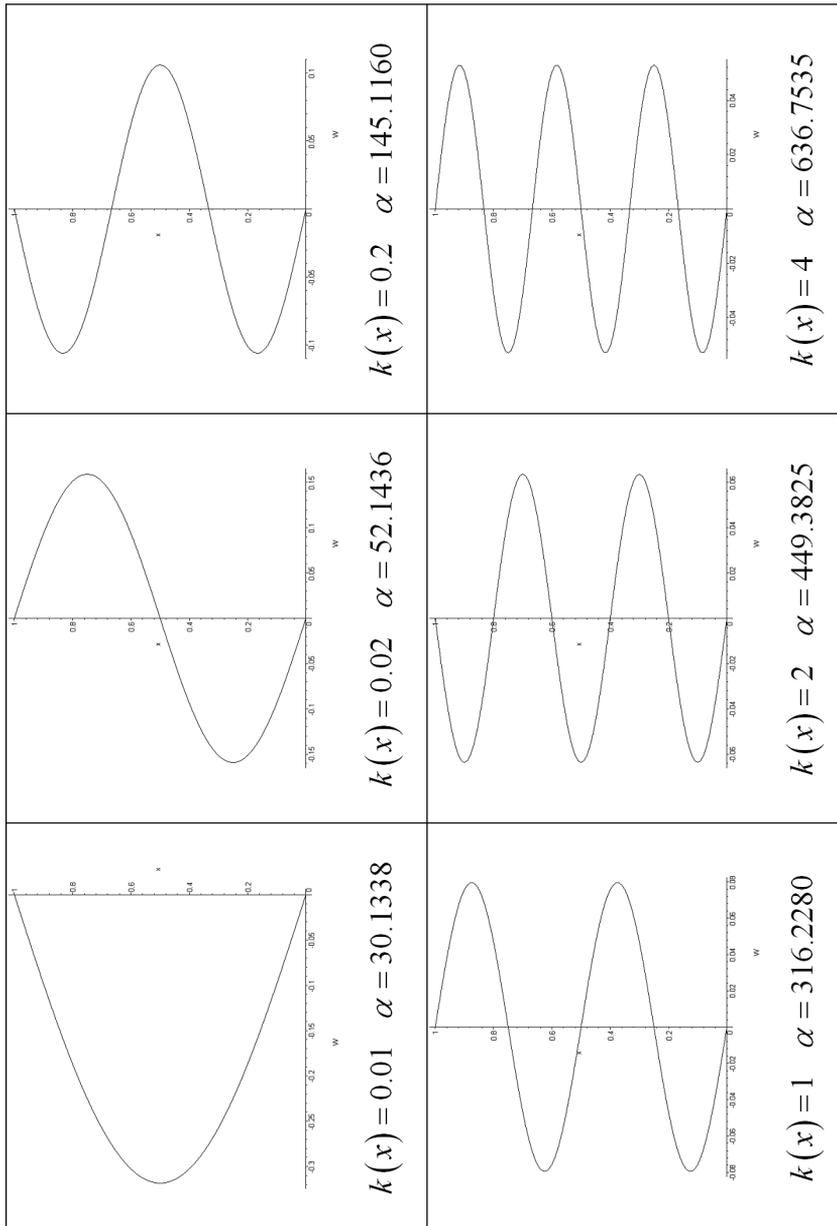


Figura 4. Primer modo normal para diferentes valores $k(x)$ y $N = 20$.

De este incremento en el número de semi-ondas del primer modo surge una pregunta: ¿A partir de qué valor del resorte pasa de una sola semi-onda a dos semi-ondas? Este valor existe y es C tal que $0.01 < C < 0.02$ (como puede observarse en la figura 4). La solución del problema para $k(x) = C$ presenta un auto valor de multiplicidad 2, es decir $\alpha_1 = \alpha_2 = 49.348022$, cuyas formas modales son semejantes al primer y segundo gráfico de la figura 4, es decir, la columna puede fallar con cualquiera de las dos formas modales.

Otro aspecto importante surge para el caso de una columna con momento de inercia constante, tal como se muestra en la Tabla 3. En ella se aprecia que con el incremento en una cierta cantidad del valor de k , la forma modal que presenta una sola semi-onda (el primer gráfico de la Figura 1), se desplaza al segundo modo, tercer modo y así sucesivamente, permutando la posición con la forma modal inmediatamente posterior.

Tabla 3: Modos con una sola semi-onda.

$k(x)$	i (modo con una sola semi-onda)	α_1	α_i
0.010	1	35.199900	35.199900
0.025	2	55.309852	73.195344
0.050	3	71.141287	136.521084
0.070	4	83.806435	187.181676
0.100	5	102.804157	263.172564
0.150	6	131.043599	389.824043
0.200	7	145.115986	516.475523

5. Conclusiones

El método de Galerkin se presenta simple de implementar en computadoras, y sin perder flexibilidad para el análisis de diferentes problemas, puede emplearse incluso en aquellos de interés en la ingeniería aplicada. Para los casos más comunes tales como los mostrados en la bibliografía, se observa una rápida convergencia a los valores exactos con el incremento del número de funciones en la base del espacio V^h . Durante la resolución de problemas con apoyo elástico continuo ($k(x) \neq 0$), se observó una disminución en la velocidad de convergencia, por lo que se tuvo que recurrir a bases de 40 funciones para obtener valores de α_1 con una precisión de 4 dígitos significativos. El fenómeno del autovalor de multiplicidad 2 con dos formas modales diferentes, también se presenta en problemas de vibraciones naturales, tal como fue analizado por Albarracín et al. (2004). La practicidad del método de Galerkin posibilita la resolución de problemas de ingeniería aplicada, es por ello que los autores prevén en próximos trabajos, publicar el análisis de casos particulares.

6. Agradecimientos

El autor agradece el apoyo y financiamiento del Consejo de Investigaciones de la UCASAL para la realización del presente trabajo (Proyecto 115/13).

7. Referencias

Albarracín, C.M., Zannier, L. y Grossi, R.O.(2004). «Some observations in the dynamics of beams with intermediate supports». *Journal of Sound and Vibration*, vol. 271, 475-480, Academic Press.

Atkinson, K. y Han, W. (2001). «Theoretical Numerical Analysis. A Functional Analysis Framework». Springer-Verlag, New York.

Rektorys, K.(1980). «Variational Methods in Mathematics, Science and Engineering».D. Reidel Co.

Wang C.M., Wang, C.Y. y Reddy, J.N. (2005). «Exact Solutions for Buckling of Structural Members». CRC Press.

Análisis de las condiciones locales para introducir transformadores de distribución de núcleos amorfos¹

**Soledad Gamarra² - Jorge Horacio Giubergia³ - José Alderete³,
Manuel Luis Zambrano² - Javier Moya²**

Resumen

Los transformadores de distribución de núcleo amorfo (AMDT por sus siglas en inglés) son, desde el punto de vista energético, los transformadores más eficientes que existen en el mercado. Su eficiencia radica en que presentan una reducción de las pérdidas en el núcleo de un 70% respecto de las de los transformadores tradicionales. Esta tecnología se viene adoptando de una manera creciente y sostenida desde 1980, con fábricas instaladas no sólo en países altamente desarrollados, como Estados Unidos, Japón, China y de la Unión Europea, sino también en países en desarrollo como India (con la mayor cantidad de AMDT del mundo) y, más recientemente, en Brasil.

A pesar de los comprobados beneficios, tanto a nivel energético como ambiental, los AMDT todavía no son empleados a nivel mundial en forma masiva por las empresas distribuidoras de energía eléctrica. En lo que respecta a nuestro país, si bien varios grupos de investigación vienen avanzando en el estudio de estos materiales, aún falta que se generen políticas que favorezcan el desarrollo del mercado necesario para su implementación, como ser, promoción de proveedores, concientización de consumidores y promulgación de la legislación pertinente. En este trabajo se realiza una breve descripción de los AMDT, exponiendo sus ventajas. Se analizan las principales razones que limitan su aplicación, con el objetivo de proponer algunas acciones tendientes a la incorporación de esta tecnología, relacionadas con la difusión del conocimiento y con un marco legal que incentive la producción de los AMDT en empresas locales, que se incluya dentro de las políticas de ahorro energético vigentes.

Palabras clave: transformadores de alta eficiencia - metales amorfos - transferencia y apropiación de tecnología - Políticas de Eficiencia Energética

Abstract

Amorphous metal distribution transformers (AMDT) are the most energy efficient transformers on the market. Their efficiency is based on reducing core losses, a 70% compared with traditional

¹ Este trabajo fue presentado en VIII° Congreso de Ingeniería Industrial (COINI) – Córdoba Argentina – Noviembre 2015.

² GIM-IESIING, Universidad Católica de Salta, INTECIN UBA-CONICET.

transformers. This technology is being adopted in a growing and steadily manner since 1980, with production plants not only in highly developed countries like the United States, Japan, China and from the European Union but also in developing countries like India (with the largest number of AMDT in the world) and more recently, in Brazil.

Despite the proven benefits, in both energy and environmental level, electricity distribution companies worldwide still not employ AMDT on a massive scale. With regard to our country, although several research groups have advanced in the study of these materials, policies that promote market development for their implementation are still to be generated, such as, promotion of suppliers, consumer awareness and enactment of relevant legislation.

This paper presents a brief description of AMDT, exposing their advantages. We analyze the main reasons that limit their application in order to propose some actions to the incorporation of this technology, concerning the dissemination of knowledge and a legal framework that encourages the production of AMDT in local businesses, to be included in energy-saving existing policies.

Keywords: high efficiency transformers - Amorphous Metals - Technology Transfer and Ownership - Energy Efficiency Policies

1. Introducción

En los últimos años ha aumentado el interés por incrementar la eficiencia en los sistemas de distribución de energía, con el objetivo de disminuir las pérdidas, mejorar la oferta energética y disminuir las emisiones de gases de efecto invernadero (GEI). Las pérdidas en los transformadores de distribución constituyen alrededor de un tercio de las pérdidas totales en el sistema de transmisión y distribución de energía (Luciano y de Castro, 2011), por lo que su disminución permitiría obtener importantes mejoras en la eficiencia de todo el sistema de distribución eléctrica.

En este sentido, los transformadores de núcleo amorfo disminuyen significativamente las pérdidas en relación a los transformadores utilizados tradicionalmente, de hierro silicio de Grano Orientado (FeSi GO). Esto se debe a que las chapas de metal amorfo con las que se desarrolla su núcleo, presentan un campo coercitivo mucho menor al FeSi GO, 2,6 A/m y 26 A/m respectivamente (Kubota, et al., 2011) y una resistencia eléctrica de un orden de magnitud mayor que la de un metal cristalino. La combinación de ambas propiedades genera una disminución de las pérdidas en el núcleo de aproximadamente el 75% [3] con respecto al material tradicional.

El primer transformador de núcleo amorfo (AMDT, por sus siglas en inglés, *Amorphous Metal Distribution Transformers*) fue desarrollado por General Electric y puesto a prueba en EEUU en 1979 (De Cristofaro, 1998). La disminución de las pérdidas energéticas fue significativa, pero las cintas del núcleo presentaban menores valores de inducción magnética, por lo que el transformador resultaba más grande y más pesado. El aumento en el costo de la energía y de los materiales tradicionales y la obtención de metales amorfos con mayor inducción magnética permitió que países como Estados Unidos y Japón continuaran trabajando para obtener AMDT más competitivos.

No pasó mucho tiempo hasta que la empresa Hitachi Metals, de Japón, se pusiera a la cabeza de la fabricación de chapas de material amorfo para los núcleos de los transformadores,

alcanzando una producción de 60.000 ton anuales en 2015 en este país (Radulovic, 2015) y de 45.000 ton anuales en su planta de Estados Unidos. India es el país con mayor cantidad de AMDT del mundo, con una capacidad de producción de chapas de más de 20 Ton mensual, y en los últimos años, la empresa China Advanced Technology & Materials ha alcanzado una capacidad de producción de 40.000 ton en el año (De Almeida et al., 2013)..

En cuanto a los AMDT en sí, en el año 2013 Estados Unidos poseía un 10% de estos transformadores en el mercado, mientras Japón alcanzaba un 30% (De Almeida et al., 2013); Hitachi Metals sigue siendo considerada la fábrica más grande a nivel mundial, con unas 150.000 unidades de AMDT vendidas hasta 2011, durante los primeros 20 años de su ingreso al mercado (Inagaki et al., 2011) (a países de Europa pero también a China y a India). Más recientemente, en el año 2012, ABB fue la primera gran empresa europea en incorporar la tecnología del metal amorfo para la fabricación de sus transformadores. Estos nuevos transformadores constituyen una importante herramienta para alcanzar los requerimientos de la Unión Europea que impone límites en el consumo energético a los países miembros a partir de 2020. Otros países que disponen de fábricas de AMDT son India, China, Taiwan, Bangladesh, Corea, Filipinas, Tailandia y Australia (Luciano y de Castro, 2011).

Un caso más cercano a Argentina es el de Brasil, donde los estudios académicos iniciaron prácticamente en el mismo tiempo que en nuestro país, en la década del 1990. Sin embargo, la transferencia a la industria fue más agresiva y hacia el año 2010 se habían desarrollado al menos cinco fábricas de AMDT (Luciano y de Castro, 2011).

Todo lo mencionado demuestra que la tendencia a utilizar estos transformadores se encuentra en aumento. La ampliación de la producción a nivel mundial y el incremento de la competencia traerán, como consecuencia, una disminución en su precio y una mayor necesidad de estudios para la producción de nuevas aleaciones y productos (un ejemplo son los motores eléctricos altamente eficientes que se comenzaron a producir con estos materiales).

El claro desarrollo en la producción mundial de este tipo de transformadores está, por supuesto, acompañado por la adopción de esta nueva tecnología por parte de las empresas de distribución (y por particulares). Sin embargo, su incorporación no es aún representativa de los potenciales beneficios que se les confiere a los AMDT.

2. Pérdidas de un transformador

Para poder comprender el beneficio de los transformadores de núcleo amorfo debemos entender cuáles son las pérdidas de un transformador tradicional y su régimen de cargas.

Las pérdidas que se producen en un transformador pueden dividirse principalmente en tres tipos: pérdidas por histéresis, pérdidas por corrientes de Foucault (o corrientes parásitas) y pérdidas en el bobinado de cobre. Las pérdidas por histéresis y por corrientes parásitas se producen en el núcleo del transformador, por lo que también se las llama pérdidas en el hierro.

Cuando el transformador se encuentra en vacío, es decir, el bobinado del secundario es un circuito abierto, las pérdidas en el cobre del bobinado primario pueden despreciarse frente a las pérdidas en el hierro. Por lo tanto, las pérdidas en vacío se consideran compuestas fundamentalmente por las pérdidas en el hierro (Georgilakis, 2011), independientes de la carga del transformador y presentes siempre que este se encuentre energizado.

Por otra parte, cuando el transformador se encuentra con carga, las pérdidas en el bobinado de cobre se vuelven más significativas mientras mayor sea el porcentaje de carga y se deben a las pérdidas en forma de calor debido a la corriente que circula por el conductor.

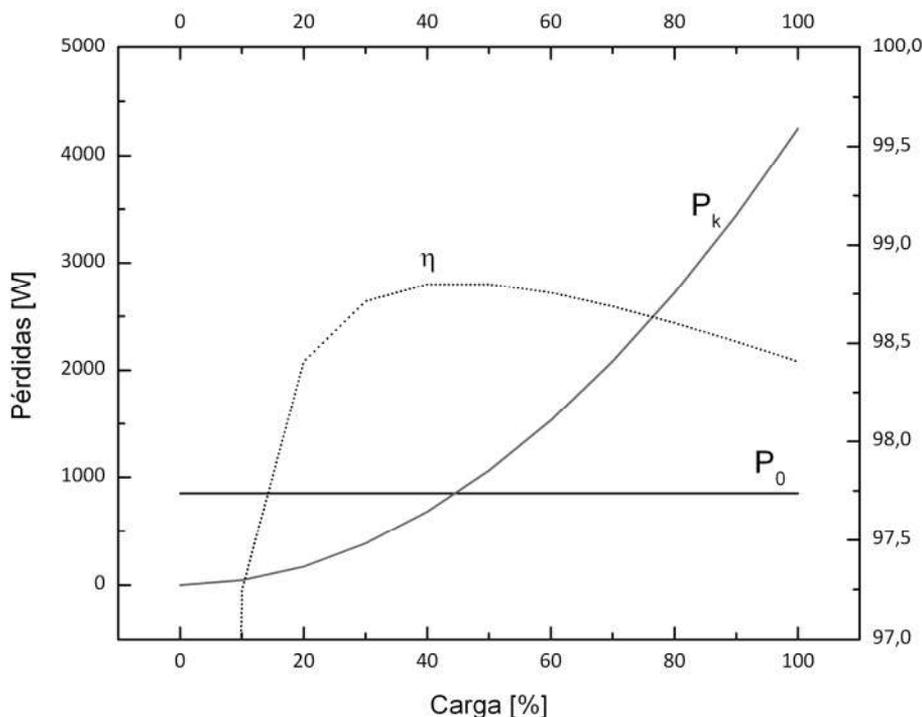


Figura 1. Pérdidas máximas y eficiencia de un transformador de 315 kVA conectado a 13.2 kV.

En la Figura 1 se presentan las pérdidas de un transformador de 315 kVA conectado a 13,2 kV y la eficiencia correspondiente (calculada a partir de las pérdidas) en relación al porcentaje de carga. Se toman como referencia los valores máximos admisibles en Argentina, establecidos por la norma IRAM 2250. Puede observarse que la eficiencia máxima corresponde a un porcentaje de carga de aproximadamente 45 %.

Como se mencionó previamente, el reemplazo del núcleo de FeSi GO en un transformador por uno de estructura amorfa genera una disminución de alrededor del 75% de las pérdidas en el hierro [3]. Las pérdidas se pueden reducir entre un 37 y un 45% si el transformador está cargado al 40% o un 13% si el transformador estuviera a plena carga (100%) (Getson, 2014).

En la Figura 2 se muestra la evolución con el aumento de la cantidad de flujo magnético de las pérdidas magnéticas del material tradicional (Urata, 2011) y de un material amorfo desarrollado a partir de materias primas de calidad comercial (Gamarra et al., 2014) a distinta frecuencia. Puede observarse que la aleación amorfa presenta pérdidas significativamente menores, alcanzando valores hasta un 70% menores al material tradicional y con un valor 40% menor correspondiente a la imanación máxima del amorfo.

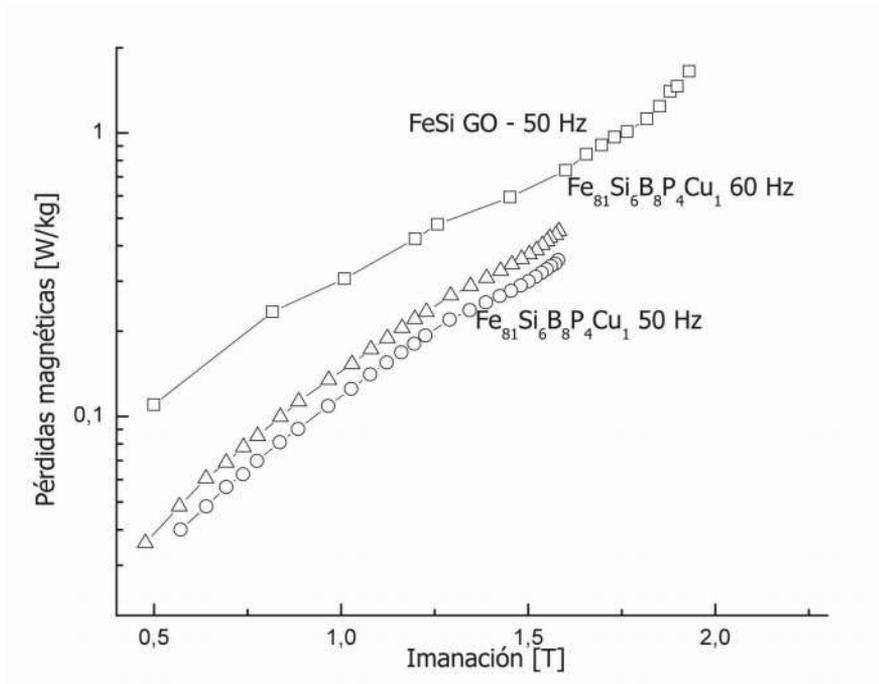


Figura 2. Pérdidas en función del flujo magnético del FeSi GO y de un metal amorfo.

3. Pérdidas de distribución en la Argentina

La falta de un sistema preciso de medición hace que la evaluación de las pérdidas de distribución resulte ambigua en algunos aspectos. Las pérdidas totales son calculadas como la diferencia entre el volumen de energía facturado (comprado por la distribuidora) y el volumen de energía vendido en un período de tiempo y se dividen en pérdidas técnicas y pérdida no técnicas. Las pérdidas técnicas están relacionadas con la energía que se pierde en el transporte y la distribución dentro de la red, y esa energía se manifiesta principalmente en forma de calor (calentamiento de los conductores eléctricos) y a veces en forma sonora (el zumbido típico de algunos transformadores y redes de alta tensión). En las pérdidas no técnicas están consideradas todos los demás factores no facturados, como el hurto de energía, errores administrativos y técnicos, etc. Según el informe del sector energético nacional del año 2005, las pérdidas totales en ese entonces resultaban ser muy dispares de acuerdo a cada provincia, desde un 4,77% hasta un 37,64% (ver Tabla 1). La gran discrepancia ocurriría principalmente debido a las pérdidas no técnicas.

Tabla 1. Pérdidas totales año 2005

Provincia	Pérdidas	Provincia	Pérdidas
Capital Federal y GBA	12,17%	Mendoza	6,67%
Buenos Aires	7,70%	Misiones	24,77%
Catamarca	11,58%	Neuquén	6,91%
Chaco	23,02%	Río Negro	4,94%
Chubut	6,49%	Salta	11,42%
Córdoba	28,47%	San Juan	8,26%
Corrientes	37,64%	San Luis	11,57%
Entre Ríos	15,68%	Santa Cruz	4,77%
Formosa	33,98%	Santa Fe	22,60%
Jujuy	8,01%	Santiago del Estero	17,64%
La pampa	18,08%	Tierra del Fuego	11,03%
La Rioja	9,79%	Tucumán	9,16%
		Total	13,66%

La empresa distribuidora EDENOR ha realizado una estimación del valor de las pérdidas técnicas por comparación con países que disponen de redes eléctricas similares, obteniendo un valor de alrededor del 7% (mientras que países como Holanda, Alemania y Bélgica presentan valores del orden del 4%). La regulación argentina ha fijado un valor de pérdidas aceptables para las empresas distribuidoras (EEDD) del 10 % (promedio) de manera tal que las pérdidas reportadas mayores a ese valor deben ser asumidas por las EEDD. En Salta, el Ente Regulador de los Servicios Públicos en su Resolución N° 833/12 establece los factores de pérdida para pasar a los costos de energía y potencia en cada nivel de tensión, los cuales se muestran en la Tabla 2.

Tabla 2. Factores de Pérdidas reconocidos por Res EnReSP 833/12

Pérdidas	Energía	Potencia
Alta Tensión (AT)	0,09 %	0,12 %
Media Tensión (MT)	2,35 %	2,88 %
Baja Tensión (BT)	9,33 %	10,50 %

Es fácil deducir que las EEDD priorizarán reducir las pérdidas no técnicas «implementando tecnologías especiales y de supervisión, tales como redes considerablemente más altas a las que no puede accederse utilizando escaleras comunes, defensas cerca de los postes de luz, cables concéntricos, medidores reforzados y corte del servicio de electricidad, entre otros recursos» (Edenor, 2015) que las técnicas, dado que las medidas de reducción de pérdidas no técnicas tienen un alto impacto en la economía de las EEDD.

4. Barreras para la adopción de los amdt

El costo de un AMDT ha sido, inicialmente, una de las mayores barreras. Si bien, en su inicio, su costo con respecto a los transformadores tradicionales a igualdad de prestaciones resultaba ser un 30% mayor, en la actualidad este porcentaje se ha reducido prácticamente en una a tercera parte, resultando ahora del orden del 23%, gracias al incremento en la producción de las chapas, tras la caída de la protección intelectual (patentes) y la entrada en el mercado de otras fábricas productoras de este material. Por ejemplo, el precio FOB de un transformador AMDT hecho en China de 315 kVA, en 13,2 kV, es de U\$ 7.385, mientras que un transformador tradicional de iguales características en Argentina cuesta aproximadamente \$ 57.000 (U\$ 6.200). A pesar del costo extra, el precio queda amortizado mucho antes de los 10 años de vida (De Cristofaro, 1998; Georgilakis, 2011; Getson, 2014) y, si tenemos en cuenta además, que la vida útil de un transformador es de alrededor de 50 años, el beneficio económico es mucho mayor.

Es importante mencionar que el precio de la tecnología tiene una tendencia decreciente (por incorporación de nuevos fabricantes, mayores volúmenes de producción, difusión de la tecnología en sí misma), mientras que el costo de la energía eléctrica es siempre creciente. Esto permite estimar que en un futuro no muy lejano, el costo de los transformadores AMDT será similar al de los transformadores tradicionales, por lo que la cuestión económica no será la única consideración en el momento de elegir una u otra tecnología.

Dejando aparte el costo de adquisición, a pesar del comprobado beneficio en el ahorro energético y en el cuidado del ambiente, el uso de los AMDT aún no se ha generalizado a nivel mundial y en nuestro país no existen empresas que produzcan ni utilicen estos equipos. La adopción de transformadores más eficientes nunca fue fácil ni por parte de los gobiernos ni por parte de las EEDD (McDermott et al., 2000). Entre las principales razones por las que su aplicación aún no es masiva, se pueden mencionar las siguientes:

1. Las compañías distribuidoras de Energía Eléctrica, dependiendo de la eficiencia de su infraestructura, no muestran mucho interés en disminuir las pérdidas técnicas de la red de distribución de energía ya que las inversiones no cuentan con un rápido período de recuperpo.
2. Las pérdidas totales de la distribuidora son socializadas en las facturas de todos los usuarios (se admite que un porcentaje se pase a los costos).
3. Falta de transferencia del KNOW HOW a las posibles pymes que podrían abastecer de AMDT al mercado.
4. Falta del desarrollo del mercado tanto de proveedores como de consumidores de AMDT.
5. Falta de un marco regulatorio que fomente el uso de los AMDT.
6. Ignorancia de las ventajas y existencia de estos transformadores.

5. Acciones propuestas para promover el empleo de AMDT

De acuerdo a las barreras identificadas en el punto anterior, se propone actuar tanto en la normativa que regula el sistema de distribución eléctrica, como en la transferencia de tecnología a las empresas productoras de transformadores, a fin de lograr mejoras en el sistema de distribución de energía que impliquen un ahorro energético y de divisas y, en consecuencia, lo vuelvan más amigable con el ambiente.

5.1. Cambio en la regulación y propuesta de un plan de mejora de la eficiencia de los transformadores

Si bien en los contratos de concesión de las empresas distribuidoras se establece la necesidad de una revisión quinquenal del nivel de pérdida totales admisibles para la distribuidora, sería importante el aumento de la exigencia en la eficiencia de la distribución (esto ocurrió en la primera revisión quinquenal en Salta), disminuyendo los porcentajes de pérdidas totales admitidos. Al mismo tiempo, se requiere un marco regulatorio que exija la incorporación gradual de los AMDT a partir de un determinado año, en el cual se prevea que ya estén disponibles en el mercado local y que su producción pueda satisfacer la demanda de los mismos.

5.1.1. Estándares de Eficiencia de los Transformadores

Las pérdidas máximas admisibles para los transformadores de distribución en la Argentina están fijadas por la Norma IRAM N° 2250, la cual establece los valores máximos permitidos para las pérdidas en vacío y para las pérdidas en el cobre. Normativas semejantes son utilizadas por la mayoría de los países (UE, China, Brasil, Chile). En la Tabla 3 se han comparado los valores admitidos a nivel nacional respecto de los admitidos en otros países para un transformador tipo 13.200V/380V-200V de 315 kVA de corriente alterna de 50 Hz.

Tabla 3: Comparación de eficiencias admitidas en distintos países

Pot. Nominal (kVA)	Tipo de núcleo	P0 (W)	Pk (W)	Ef (50%)	País/Norma
315	Tradicional	850	4250	98.80%	Argentina IRAM N° 2250
315	Amorfo	180	3900	99.27%	Xiamen Electric Home Science & Technology Development CO.,LTD. (China)
315	Amorfo	-	-	99.66%	ABB-AMDT (ABB REview, 2012)
315	Tradicional de alta eficiencia	520	3200	99.17%	(E-MT-009) – Enersis-Endesa Chile

Otra forma de caracterizar los transformadores consiste en indicar la eficiencia cuando se encuentra cargado al 50% de su carga nominal, como indica la Ecuación (1).

$$\eta(50\%) = \frac{P_{nom}(50\%)}{P_{nom}(50\%) + P_0 + P_k(50\%)} \quad (1)$$

En donde $P_{nom}(50\%)$ es la mitad de la potencia nominal de diseño del transformador, P_0 las pérdidas en vacío y $P_k(50\%)$ las pérdidas del cobre evaluadas al 50% (que serán $\frac{1}{4}$ de las P_k a la potencia nominal, ver Figura 1). Se utiliza el valor de $P_{nom}(50\%)$ dado que alrededor de ese porcentual de carga los transformadores encuentran su mayor eficiencia, como se observa en la Figura 1. Este método es empleado en EEUU, Canadá, Australia, Nueva Zelanda, Corea del Sur, mientras que en India se utiliza una combinación de los dos métodos (pérdidas máximas y eficiencia al 50% de carga). En México se utiliza un método similar, pero considerando la carga al 80%.

Tomando los datos de pérdidas máximas permitidas para distintos transformadores trifásicos según las normas IRAM 2250 y transformándolos en Eficiencia al 50% obtenemos la curva de la Figura 3. En la misma gráfica, se han adjuntado las eficiencias actuales de Estados Unidos, de la Unión Europea (programas C-C') y las propuestas para la UE en el programa EN50464; los valores propuestos en Australia, así como las previstas por el programa Top Runner de Japón (Blackburn, 2007). Se puede observar que, en comparación, nuestra normativa se encuentra lejos de los estándares actuales de los países desarrollados.

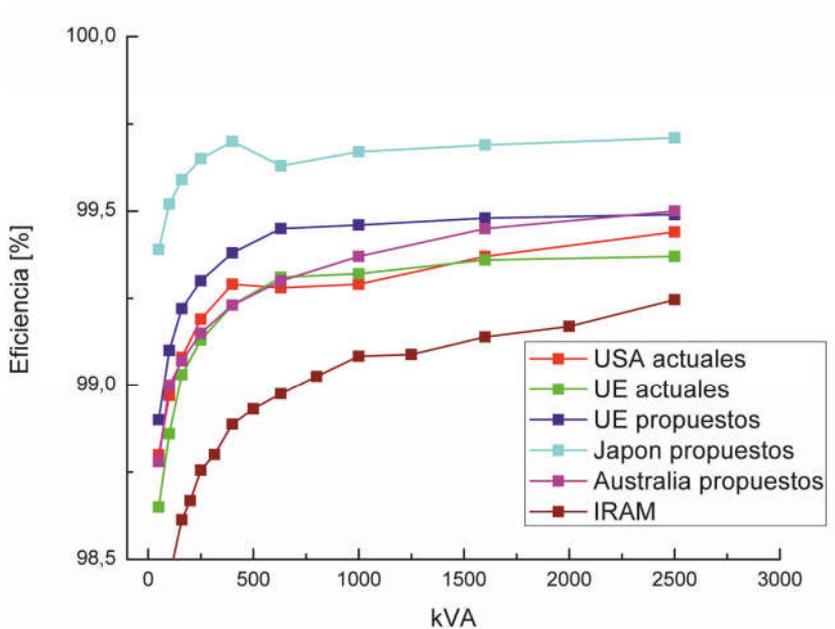


Figura 3. Eficiencias mínimas para diferentes países.

En 2007, el gobierno de Australia promovió el programa para incrementar la eficiencia de sus transformadores en un 0,25 - 0,08% para transformadores de 25 a 2500 kVA respectivamente, en relación a los valores actuales. Por ejemplo, para el caso del transformador de 315 kVA, que venimos considerando, se prevé una mejora en la eficiencia de 0,15%, pasando del valor actual de 99,04 al propuesto del 99,19%. Es importante notar el interés en incrementar una cantidad aparentemente despreciable de la eficiencia. Como se puede apreciar en la Tabla 3, estos estándares son fácilmente superados por los transformadores amorfos.

5.1.2 Pérdidas vs. Eficiencia

Otra normativa a reconsiderar se relaciona con el método de caracterización de las exigencias mínimas en las propiedades eléctricas de un transformador: cuantificar las pérdidas, como se realiza en Argentina hasta el momento, o establecer la eficiencia mínima. Un ejemplo práctico para confrontar estas dos normas es el siguiente: en el 2009, una EEDD de la región (ABB, 2012) detalló las especificaciones técnicas que deberían cumplir los transformadores adquiridos por dicho grupo para ser instalados en algunas redes de distribución en Latinoamérica, estableciendo tres niveles de pérdidas para los transformadores. En el nivel III, el de los más eficientes, se establece un valor máximo para las pérdidas en vacío de 520 W y para las de plena carga de 3200 W, para un transformador de 315 kVA. Si se comparan estas exigencias con los valores del transformador amorfo expuesto en la Tabla 2, observamos que el AMDT no cumple con los requisitos, ya que las pérdidas a plena carga (3900 W) son mayores que las especificadas y por lo tanto quedaría fuera de este nivel. Sin embargo, si se considera el parámetro de eficiencia, se observa que los dos AMDTs que se encuentran en la misma tabla superan cómodamente las especificaciones del transformador tradicional de alta eficiencia. De aquí surge que los estándares, tales como las Normas IRAM 2250, pueden quedar desactualizadas al no tener en cuenta la nueva tecnología de los núcleos amorfos con unas pérdidas en vacío extremadamente bajas, por lo que se debería actualizar la normativa a fin de incluirlos adecuadamente.

5.1.3 Regulación para el empleo de AMDT

En Argentina existen antecedentes de regulaciones que lograron que ciertas políticas de mejoras avancen paulatinamente, involucrando y siendo recibidas por toda la sociedad.

Un ejemplo es el caso de la eliminación de PCBs (policlorobifenilos, policloroterfenilos, etc) cuya ley fue promulgada en noviembre de 2002. En la Ley de Presupuestos mínimos para la gestión y eliminación de PCBs (Ley 25.670) y en el Decreto 853/2007, que la reglamenta, se establece la creación de un Consejo Técnico integrado por representantes de distintas entidades interesadas, para asesoramiento y asistencia en temas relacionados con la gestión de los PCBs. Asimismo, establece el desarrollo de guías técnicas y de estrategias para la difusión pública de información y participación ciudadana. Finalmente establece un plazo de 7 años para alcanzar la descontaminación total y la eliminación de estas sustancias.

Este ejemplo permite pensar que es posible regular la incorporación gradual de estos transformadores, avanzando hacia el reemplazo total de los transformadores tradicionales, a partir de un determinado año, en el cual se prevea que ya estén disponibles en el mercado local y que su producción pueda satisfacer la demanda de los mismos.

5.2. Transferencia tecnológica

Es necesario establecer una cooperación activa de transferencia de tecnología y know how entre los institutos de investigación y las pymes interesadas en producir los AMDTs.

En la Argentina, el estudio de los metales amorfos para uso magnético comenzó en la década del 1990 con varios proyectos de investigación aplicada en el tema. Hacia el 2010 el Ministerio de Ciencia y Tecnología (MINCYT) financió un proyecto de transferencia tecnológica aún en vigencia (FONARSEC FS-NANO03) que contempla la producción de las chapas amorfas en la Argentina. Otro proyecto de transferencia es el denominado FITS 2013 UREE 0003 de «Implementación de Redes Eléctricas Inteligentes» también financiado por MINCYT en la provincia de Salta, que prevé la incorporación de unos 15 transformadores de núcleo amorfo a modo de red piloto.

Por otro parte, en el mercado internacional, se encuentra disponible equipamiento específico para efectuar distintas etapas del proceso productivo de los AMDT (corte de chapas de metal amorfo, tratamientos térmicos, etc).

6. Conclusiones

En la actualidad, los transformadores de núcleo amorfo son los más eficientes del mercado internacional, que muestra cada vez más aceptación y una producción mundial que se incrementa año a año.

Si bien el costo de adquisición es aún alrededor de un 20% mayor que el de los transformadores tradicionales, por otro lado, el aumento de la producción a nivel mundial, el vencimiento de los registros de propiedad intelectual, la posibilidad de producción local de la chapa amorfa, el encarecimiento de la energía eléctrica y la necesidad de mejorar la eficiencia del sistema de distribución a fin de reducir la emisión de gases de efecto invernadero, indican que es factible la implementación masiva de los AMDTs.

Las normas argentinas sobre los valores admisibles de pérdidas en los transformadores de distribución se encuentran desactualizadas en relación a los valores establecidos en las regulaciones de numerosos países, a pesar de que la tecnología actual permite incrementar sensiblemente la eficiencia. Asimismo, la caracterización de los transformadores en base a su eficiencia, permitiría valorizar la importante disminución en las pérdidas en vacío que se produce en los AMDT, incentivando su utilización por parte de las empresas de distribución.

7. Referencias

ABB Review - The corporatetechnical journal, 2/12 (2012)

Blackburn, T. R. (2007). Technical Report - Distribution Transformers: Proposal to Increase MEPS Levels. Australia.

De Almeida, Aníbal T.; Martins, Fernando; Santos, Bruno. (2013). «Ecodesign Requirements for Power, Distribution and Small Transformers». Impact Assessment. University of Coimbra, Portugal.

- DeCristofaro, Nicholas. (1998). «Amorphous Metals in Electric-Power Distribution Applications». Materials Research Society. Vol. 23, 5, 50 - 56.
- «Edenor - RI - Pérdidas de Energía.» Disponible en: http://www.edenor.com.ar/cms/SP/EMP/RI/EST_perdidas.html. [Accedido: 02-sep-2015].
- Especificación técnica: transformadores de distribución convencionales para instalación aérea o apoyada a nivel de superficie (E-MT-009) (2009). Enersis-Endesa. Chile.
- Gamarra, Soledad; Marta, Leonardo; Berejnoi, Carlos; Moya, Javier. (2014). «Estudio de Propiedades Magnéticas de la Aleación Fe81Si6B8P4Cu1 elaborada con materiales de baja pureza». Congreso Argentino de Ingeniería. San Miguel de Tucumán, Argentina
- Getson, Douglas. (2014). Transformadores Inmersos en Aceite - Eficiencia en redes de distribución. ABB Group. Disponible en: <https://abblibrary.s3.amazonaws.com/public/2ea6006b0851d50185257cdf004bbbc1/9AKK106103A5740%20v2%20-%20Transformadores%20inmersos%20en%20aceite%20-%20Eficiencia%20en%20redes%20de%20distribucion.pdf>
- Georgilakis, Pavlos S. (2011). «Environmental cost of distribution transformer losses». Applied Energy. Vol. 88, 3146–3155.
- Inagaki, Katsutoshi; Kuwabara, Masanao; Sato, Kohei et al. (2011). «Amorphous Transformer Contributing to Global Environmental Protection». Hitachi Review. Vol. 60, No. 5, 250-256.
- Kubota, Takeshi; Makino, Akihiro; Inoue, Akihisa.(2011). «Low core loss of Fe85Si2B8P4Cu1 nanocrystalline alloys with high Bs and B800». Journal of Alloys and Compounds. Vol. 509S, S416–S419.
- Luciano, B. A.; de Castro, W. B. (2011). «Transformadores de distribuição de energia elétrica com núcleo de metal amorfo». Revista Eletrônica de Materiais e Processos. Vol. 6.1, 01-05.
- M J McDermott et al. (2000). «Energy-efficient distribution transformers -utility initiatives». By International Copper Association European Copper Institute.
- Radulovic, Verena. (2015). «Comments from Hitachi Metals America Ltd. and Metglas Inc. regarding the ENERGY STAR Transformer Program». Hitachi Metals America.
- Urata, A., Matsumoto, H., Yoshida, S., Makino, A. (2011). «Fe-Si-B-P-Cu Nanocrystalline Alloy Ribbons With High Saturation Magnetic Flux Density Prepared Using Industrial Materials». IEEE Transactions on Magnetism. Vol. 47, 10, 3177-3179

Estimación de láminas de lluvias máximas diarias y sus parámetros estadísticos en el norte de Argentina

**Nicolás F. Guillén^{2,3}- M. Cecilia Botelli¹- Carlos M. García^{2,3}-
Carlos Gastón Catalini⁴**

Resumen

Actualmente, una gran extensión del territorio argentino no cuenta con herramientas normalizadas para predecir los parámetros de diseño hidrológico de las medidas estructurales y no estructurales requeridas para mitigar la amenaza hídrica; por lo que cada proyectista utiliza los datos disponibles y metodologías de predicción diferentes, según su grado de experiencia y criterio. Esta problemática, hace que se requiera contar con variables de diseño homogeneizados, estimados con metodologías estandarizadas de predicción disponibles en una interfaz amigable.

Este trabajo ha sido desarrollado en el marco de un esfuerzo conjunto entre el Centro de Estudios y Tecnología del Agua (CETA) de la Universidad Nacional de Córdoba, y la Universidad Católica de Salta, con el fin de estimar los valores de la lámina de lluvias máximas diarias y parámetros estadísticos en el Norte de Argentina incluyendo las provincias de Salta, Jujuy, Catamarca, Tucumán, Chaco, Formosa y Santiago del Estero.

Hasta el momento se analizaron registros de 435 estaciones pluviométricas disponibles en la región de estudio. A partir de estos registros se elaboraron las series de máximos diarios anuales. Luego se verificaron las hipótesis estadísticas básicas (presencia de datos atípicos, independencia, estacionalidad y homogeneidad) que debe cumplir una muestra para que tenga validez el análisis de estadística inferencial. Finalmente se seleccionaron registros de 367 estaciones que verificaron las hipótesis propuestas. Sobre los registros seleccionados se realizaron ajustes a 6 funciones de distribución de probabilidad para estimar los valores de máxima lámina de lluvia para los periodos de retorno entre 2 y 100 años, y los intervalos de confianza asociados a cada predicción. Se seleccionó la función Log Normal como óptima y se cuantificaron las diferencias relativas en la predicción debido a esta hipótesis. Finalmente se realizó un análisis espacial de la evolución de los valores de lámina de lluvia diaria y sus parámetros estadísticos en toda la región de estudio utilizando un Sistema de Información Geográfica.

Palabras clave: Precipitación Máxima Diaria - Norte de Argentina

¹ Universidad Católica de Salta - Grupo de Investigación de Gabinete de Hidráulica e Instalaciones Sanitarias.

² CETA. Centro de Estudios y Tecnología del Agua y Laboratorio de Hidráulica, Facultad de Ciencias Exactas, Físicas y Naturales. Universidad Nacional de Córdoba.

³ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

⁴ Grupo de Estudios Hidrológicos en Cuencas Pobremente Aforadas de la Universidad Católica de Córdoba (UCC-EHCPA).

Abstract

Currently, most of the Argentina territory has no standardized tools for predicting hydrologic design parameters for structural and non-structural management measures required to mitigate the flow risk; so the engineers use the available data and different methodologies to predict hydrologic parameters according to their experience and judgment. To achieve a properly hydrologic prediction is necessary to have homogeneous design variables and good predictions with standardized methodologies in a user-friendly interface.

This work has been developed in the framework of a joint effort between the Water Technology and Studies Center (CETA) of the National University of Cordoba and the Catholic University of Salta, in order to estimate maximum rainfall depth values and daily rainfall statistical parameters in the northern of Argentina including the provinces of Salta, Jujuy, Catamarca, Tucumán, Chaco, Formosa and Santiago del Estero.

Up to the present 435 rainfall stations were analyzed in the study region. The annual maximum daily rainfall series were calculated. Then, the basic statistical hypotheses were verified (presence of outliers, independence, seasonality and homogeneity). 367 rainfall stations verify the statistical hypotheses and 6 probability distribution functions with its respective confidence intervals were adjusted in each station in order to estimate the maximum rainfall depth values for annual recurrences between 2 and 100 years. Log Normal function was selected and the relative differences in the prediction because of this assumption were quantified. Finally, a spatial analysis of the evolution of the maximum rainfall depth values and the daily rainfall statistical parameters was developed using a Geographic Information System.

Keywords: maximum daily rainfall - north of Argentina

Introducción

El diseño hidrológico se define como la evaluación del impacto de los procesos hidrológicos y la estimación de valores de las variables relevantes para modificarlo. El diseño hidrológico se puede utilizar para mitigar amenazas hidrológicas proyectando medidas «estructurales», como una represa para el control de crecientes, sistemas de control de excedentes pluviales en obras viales, entre otras, o bien, desarrollando medidas «no estructurales» como ser la determinación de áreas de riesgo.

La escala de diseño hidrológico es el rango en magnitud de la variable de diseño dentro del cual se debe seleccionar un valor para determinar el flujo de entrada al sistema. Los factores determinantes en la selección de este valor para una obra civil son el costo y la seguridad asociada. Sobredimensionar las obras es antieconómico, mientras que sub-dimensionarlas puede ocasionar fallas con resultados catastróficos. La magnitud óptima para el diseño es aquella que equilibra criterios de costo y seguridad.

Generalmente el diseño hidrológico lo utiliza para la estimación de la Creciente de Proyecto. Sin embargo la estimación de los caudales escurridos en la creciente de proyecto está afectada por la insuficiencia estadística de los registros históricos de caudales, lo cual lleva a evaluar indirectamente estos caudales mediante el uso modelos de transformación lluvia – caudal (P-Q), los cuales son alimentados por eventos hipotéticos críticos (lluvias de diseño). Si bien tienen

otros componentes, estas lluvias de diseño parten de un vínculo esencial entre los rasgos característicos de la precipitación: la función Intensidad-Duración-Recurrencia (i-d-T). Las relaciones i-d-T y el patrón de distribución temporal de las lluvias, requeridos para estimar Crecientes de Proyecto, solo se pueden extraer de extensos registros de alta frecuencia, normalmente fajas pluviográficas. Aun reemplazando caudal por intensidad de lluvia se continúa teniendo un problema, ya que las series pluviográficas largas también son escasas, aunque menos que las hidrométricas. En cambio, es habitual otro dato de lluvia provisto por pluviómetros: la lámina de lluvia diaria. De esta manera, los técnicos y proyectistas pueden proceder a la estimación mediante un análisis de estadística inferencial de las láminas de lluvias máximas diarias asociadas a diferentes periodos de retorno, a partir de las series de datos pluviométricos registrados en diversos puntos de la cuenca.

En este trabajo se estimó la lámina de lluvia máxima diaria asociada a diferentes periodos de retorno en el norte argentino. A partir de ella, puede calcularse primero la lámina máxima asociada a una duración diferente (utilizando un coeficiente de reducción o maximización) y posteriormente es posible calcular mediante un modelo de transformación P-Q, el caudal de diseño.

Objetivos

El objetivo general del presente trabajo es la elaboración de mapas digitales de lluvias máximas diarias asociados a recurrencias de 2, 5, 10, 20, 25, 50 y 100 años para el norte de Argentina. A continuación se detalla de manera resumida la metodología empleada con la finalidad de cumplimentar el objetivo citado.

Metodología

La metodología utilizada se puede resumir de la siguiente manera:

1. Recopilación y pre-procesamiento de series pluviométricas de las distintas estaciones pluviométricas existentes.
2. Ubicación espacial de las distintas estaciones pluviométricas preseleccionadas.
3. Determinación del año hidrológico para la región de estudio.
4. Generación de la serie hidrológica de máximas lluvias diarias.
5. Verificación de hipótesis estadísticas de las series históricas de máximas lluvias diarias seleccionadas tales como independencia, estacionalidad y homogeneidad, además de la detección de datos atípicos en las series.
6. Determinación de los parámetros provenientes del análisis de estadística descriptiva (media y desvío estándar).
7. Análisis de estadística inferencial mediante el ajuste de distribución de probabilidad.
8. Interpolación de la información puntual, generación de grillas de análisis mediante el uso de un Sistema de Información Geográfica (Idrisi Selva 17.00).

1. Recopilación y pre-procesamiento de series pluviométricas de las distintas estaciones pluviométricas existentes.

Se recopiló información de la lámina de lluvia diaria precipitada, la misma se obtuvo de las estaciones pluviométricas instaladas en las diferentes provincias de la región de estudio. Para la obtención de datos se recurrió a diferentes fuentes como:

- Páginas web, en su mayoría pertenecientes a entidades gubernamentales, donde se encuentran datos históricos de lluvia diaria:

<http://bdhi.hidricosargentina.gov.ar/sitioweb/frmFiltro.aspx>

<http://dev.siia.gov.ar/series>

<http://siga2.inta.gov.ar/en/datoshistoricos/>

<http://www.tutiempo.net/clima/Argentina/AR.html>

- Instituto Nacional de Tecnología Agropecuaria (INTA).
- Secretaría de Recursos Hídricos de las diferentes provincias.

En total se analizaron 435 estaciones pluviométricas dentro del territorio. Luego se realizó una preselección, de acuerdo a cada longitud de serie de las estaciones para continuar con el análisis. El criterio de Longitud de Serie consistió, en un principio, en fijar veinte años como longitud mínima de las series anuales de lluvias máximas diarias a analizar. Este condicionante, implica rechazar puestos pluviométricos disponibles, por lo que se adopta como criterio final la aceptación de todas las estaciones con una longitud de serie mayor o igual a catorce años. Esta longitud mínima coincide con lo definido para estudios similares realizados en otras provincias de Argentina.

La Tabla 1 y la Figura 1 permiten visualizar la cantidad de estaciones que pertenecen a cada intervalo de longitud de serie, es decir, entre uno y catorce años, mayores o igual a catorce años y mayor o igual de veinte. Así, por otra parte, existen 186 estaciones con más de catorce años de registro, conformando el 53% del total; este porcentaje de estaciones son las seleccionadas para continuar con el proceso de análisis.

Tabla 1. Frecuencias de las estaciones pluviométricas en función de su longitud de serie

Longitud de serie		
Estaciones con datos	452	Porcentaje
1 < Longitud de serie < 14	138	31%
Longitud de serie ≥ 14	314	69%
Longitud de serie ≥ 20	241	53%

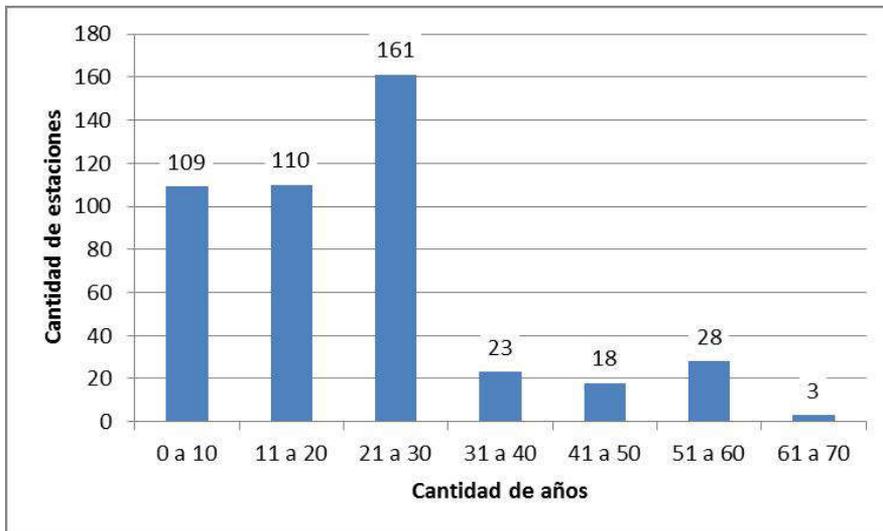


Figura 1. Longitud de registro de las estaciones pluviométricas recopiladas inicialmente.

2. Ubicación espacial de las distintas estaciones pluviométricas y sus parámetros estadísticos

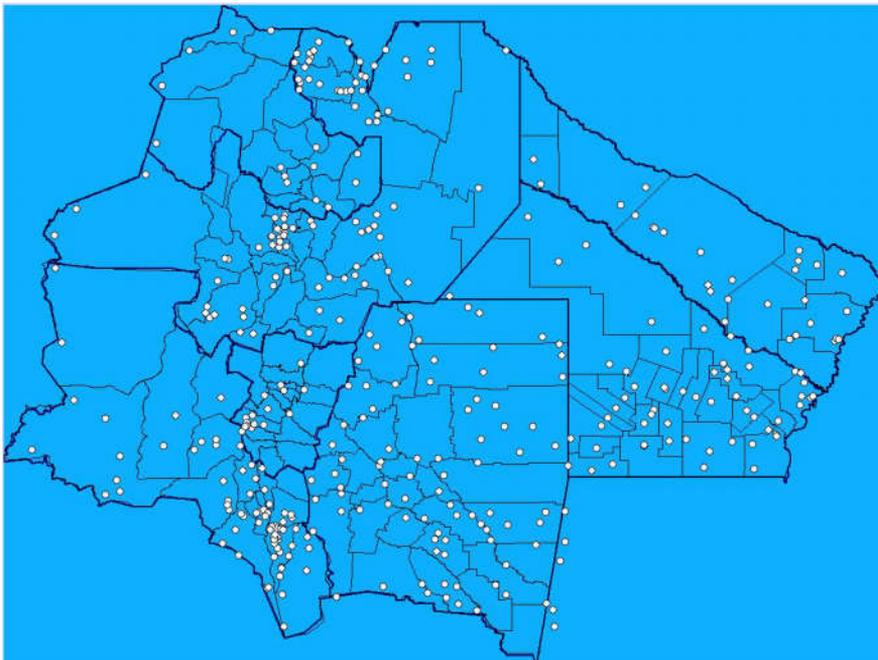


Figura 2. Disposición espacial de las estaciones pluviométricas encontradas.

En la Figura 2 puede apreciarse la disposición espacial de las estaciones pre-seleccionadas. Puede observarse una mayor concentración en la zona centro de Salta, este de Catamarca, este de Chaco y este de Formosa y en Santiago del Estero. En esta instancia del desarrollo del trabajo se incluyeron pocas estaciones en el sector de la Puna y los datos pluviométricos fueron recopilados en base a información extraída del modelo TRMM (oeste de Salta, Catamarca y Jujuy).

3. Determinación del año hidrológico para la región en estudio

Para el análisis y procesamiento estadístico de las lluvias es recomendable dividir las series históricas en «años hidrológicos». Este concepto, se define como el período anual que no fracciona el ciclo estacional de lluvia, y puede ser distinto al año calendario. Se observó, para los registros encontrados, que las máximas lluvias se daban entre los meses de octubre a abril, por ende, para el análisis se optó por determinar cómo año hidrológico al periodo desarrollado de julio a junio del año (calendario) siguiente.

4. Generación de la serie hidrológica de láminas de lluvias máximas diarias

Luego de fijar el año hidrológico se elaboraron las series hidrológicas de máximos anuales de lámina de lluvias diarias. Se pudo determinar que los valores pico de lluvia máxima diaria ocurrieron generalmente entre los meses de octubre y abril, inclusive. Por ello cuando no se contó con datos en algunos de esos meses, se prescindió de ese año hidrológico, dada la probabilidad de que en esos meses se registrara la precipitación máxima anual, lo que podría ocasionar subestimaciones en los análisis estadísticos.

5. Verificación de hipótesis estadísticas de las series históricas de máximas anuales de lluvia diaria

5.1. Detección de presencia de datos atípicos

Las series de máximos anuales de lluvia diaria de las 314 estaciones pluviométricas preseleccionadas que tenían una longitud de serie acorde fueron sometidas a la detección de datos atípicos. Si las series tenían uno o más valores atípicos, estos se removían y se realizaba el test nuevamente y si luego de ello seguían existiendo datos atípicos, la serie era descartada. En el caso que la serie verificara esta prueba, se pasaba al siguiente test. De todas las estaciones analizadas, 13 no han pasado esta prueba (Tabla 2) y fueron descartadas. Dicha eliminación representa el 4% de las estaciones que tenían una longitud de serie acorde.

5.2. Prueba de Independencia

Las 301 estaciones que verificaron la hipótesis anterior fueron sometidas a la prueba de Independencia (Wald-Wolfowitz). En la Tabla 3 se muestran las estaciones depuradas. Del total de estaciones analizadas 10 estaciones no pasaron la prueba, lo cual constituye el 3% del total de estaciones que verificaron datos atípicos.

Tabla 2. Estaciones pluviométricas que no pasaron la prueba de detección de presencia de datos atípicos

Estaciones pluviométricas
El Ceibal
Presidente Roque Sáenz Peña Aeródromo
Toropan
Ancocha
Urutau
Roversi
La Fragua
Nueva Esperanza
San Francisco
Palo Negro
Garza
Tacañitas
Aeródromo Formosa

Tabla 3. Estaciones pluviométricas que no verificaron la Prueba de Independencia de sus valores máximos diarios

Estaciones pluviométricas
Tinogasta
Tuc-Tuca
Lules – Potrero de las Tablas
Frias
Santa Catalina
Brea Pozo
Fernandez
General Pinedo
Agente Asunción Vidal Leguizamón
Villa Escobar

5.3. Prueba de Estacionalidad

Las 291 estaciones que verificaron la hipótesis anterior fueron sometidas a la prueba de estacionalidad. En la Tabla 4 se muestran las estaciones depuradas. Del total de estaciones analizadas, 8 estaciones no pasaron la prueba, lo cual constituye el 3% del total de estaciones que verificaron independencia.

Tabla 4. Estaciones Pluviométricas que no verificaron la Prueba de Estacionalidad de sus valores máximos diarios.

Estaciones pluviométricas
El Ceibal
Presidente Roque Sáenz Peña Aeródromo
Toropan
Ancocha
Urutau
Roversi
La Fragua
Nueva Esperanza
San Francisco
Palo Negro
Garza
Tacañitas
Aeródromo Formosa

5.4. Prueba de Homogeneidad a Escala Anual

Las 283 estaciones restantes fueron sometidas a la Prueba de Homogeneidad (Wilcoxon) de sus registros históricos. En la Tabla 5 se muestran las estaciones depuradas. Del total de estaciones analizadas, 4 estaciones no pasaron la prueba, lo cual constituye el 1% del total de estaciones que verificaron estacionalidad.

Tabla 5. Estaciones Pluviométricas que no verificaron la Prueba de Homogeneidad de sus valores máximos diarios.

Estaciones pluviométricas
Tinogasta
Tuc-Tuca
Lules – Potrero de las Tablas
Frias
Santa Catalina
Brea Pozo
Fernández
General Pinedo
Agente Asunción Vidal Leguizamón
Villa Escobar

Finalmente se decidió descartar la estación «Jujuy Aeródromo» debido a que esta presentaba problemas en los registros de lluvia, como ser 6 datos de lluvias que superan los 350 mm, lo cual se considera un registro excesivo. Debido a esta anomalía se decidió prescindir de la estación en cuestión (Tabla 6). Como hay varios datos dudosos (con registro excesivo), estos no son detectados por el test que detecta los datos atípicos.

Tabla 6. Estación que mostraba datos dudosos

Estaciones Pluviométricas
Jujuy Aeródromo

De esta manera se concluye que de las 452 estaciones pluviométricas de las cuales se partió este trabajo, 278 pasaron las pruebas representando estas el 61% del total. En la Figura 3 se puede observar la distribución espacial de las estaciones que pasaron las pruebas.

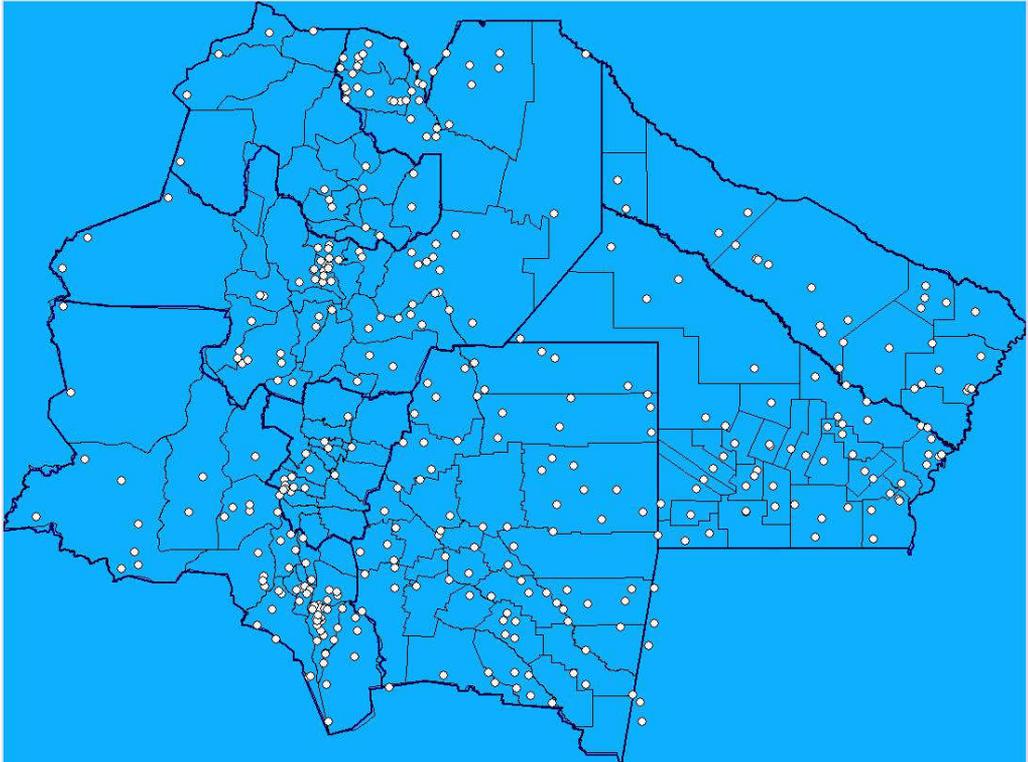


Figura 3. Estaciones finalmente seleccionadas.

6. Análisis de estadística descriptiva

De cada una de las estaciones finalmente seleccionadas, se calcularon los datos estadísticos más importantes de las series anuales de máximos de lluvia diaria: Valor máximo observado en la serie, promedio aritmético y desvío estándar para las series.

7. Análisis de estadística inferencial

En primer lugar, se determinaron las probabilidades empíricas de cada dato de lluvia diaria máxima anual observado, utilizando la Ecuación de Weibull en cada estación pluviométrica.

Luego, dichos datos se ajustaron a las 6 funciones de distribución de probabilidad teóricas siguientes:

- a) GEV (Método de Máxima Verosimilitud).
- b) Gumbel (Método de Máxima Verosimilitud).

- c) Log Normal (Método de Máxima Verosimilitud).
- d) GEV (Método de Momentos).
- e) Gumbel (Método de Momentos).
- f) Log Pearson tipo III (Método de Momentos).

Debido a que este trabajo busca extender un análisis de regionalización de lluvias máximas diarias realizado en otras regiones de Argentina (Belli et al., 2012; Caamaño Nelli y Dasso, 2003; Catalini et al., 2010; Catalini et al., 2012), es que se adopta la misma función de distribución utilizada en esos estudios: Log Normal con parámetros ajustados por el método de Máxima Verosimilitud para representar las muestras de valores de lluvias diarias máximas anuales.

Se analiza a continuación la relación existente entre la longitud de la serie de las estaciones seleccionadas con más de 14 años de datos versus la máxima diferencia porcentual entre los valores de lámina de lluvia para 100 años de periodo de retorno existente entre la función Lognormal y la función que arroja el valor de lámina más lejano a esta distribución adoptada. Se puede observar en la Figura 4 que la máxima diferencia de los valores estimados con otras funciones de distribución con respecto a la Log Normal disminuye a medida que aumenta la longitud de registro.

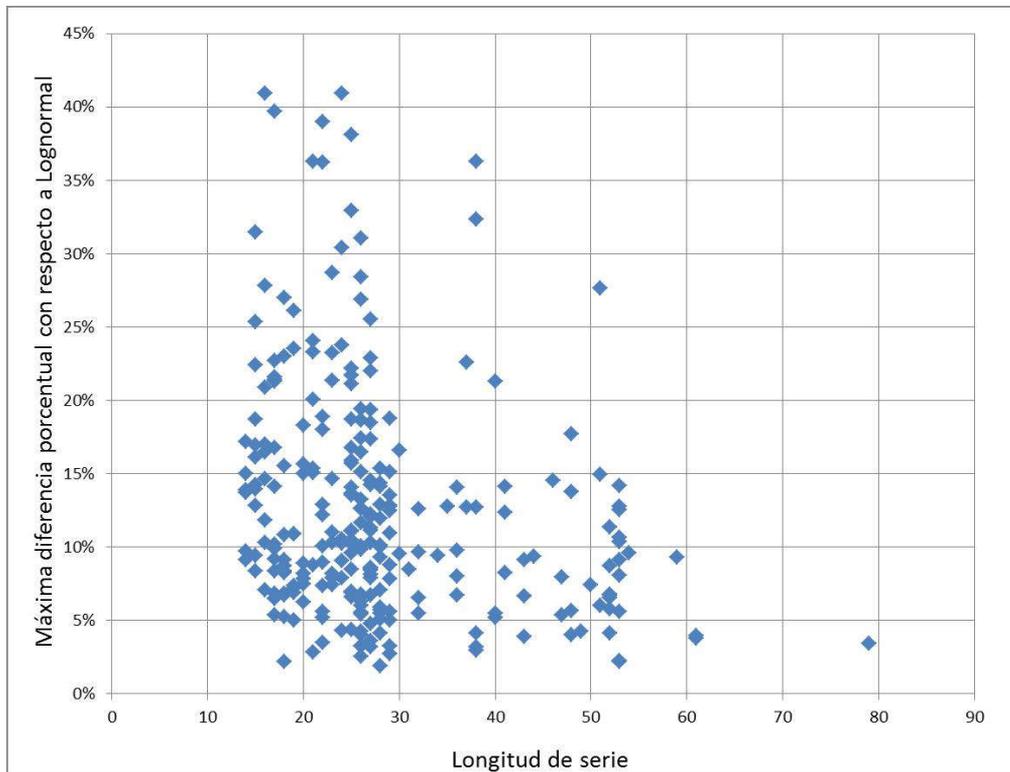


Figura 4. Evolución de la máxima diferencia porcentual entre los valores estimados con la distribución Log Normal y las otras funciones ensayadas a medida que aumenta la longitud de la serie, para la provincia de Catamarca.

A continuación se analizará la validez de adoptar la distribución Lognormal para relacionar la lámina de lluvia a diferentes recurrencias. Para analizar la importancia relativa de la selección de la función de distribución óptima se realiza un análisis para 100 años de periodo de retorno. Se elige este valor (y no los asociados a otros periodos de retorno) por ser el que más diferencia muestra en relación a los demás. En este análisis se compara la diferencia porcentual mostrada en la Figura 4 con un valor denominado «Incertidumbre del método Lognormal» representativo del tamaño de los intervalos de confianza calculado como la diferencia relativa (límite superior – valor medio) / valor medio. En la Figura 5 se puede ver el análisis realizado. Así se verifica que casi todas las funciones de distribución se hallan dentro del intervalo de confianza de la función Lognormal.

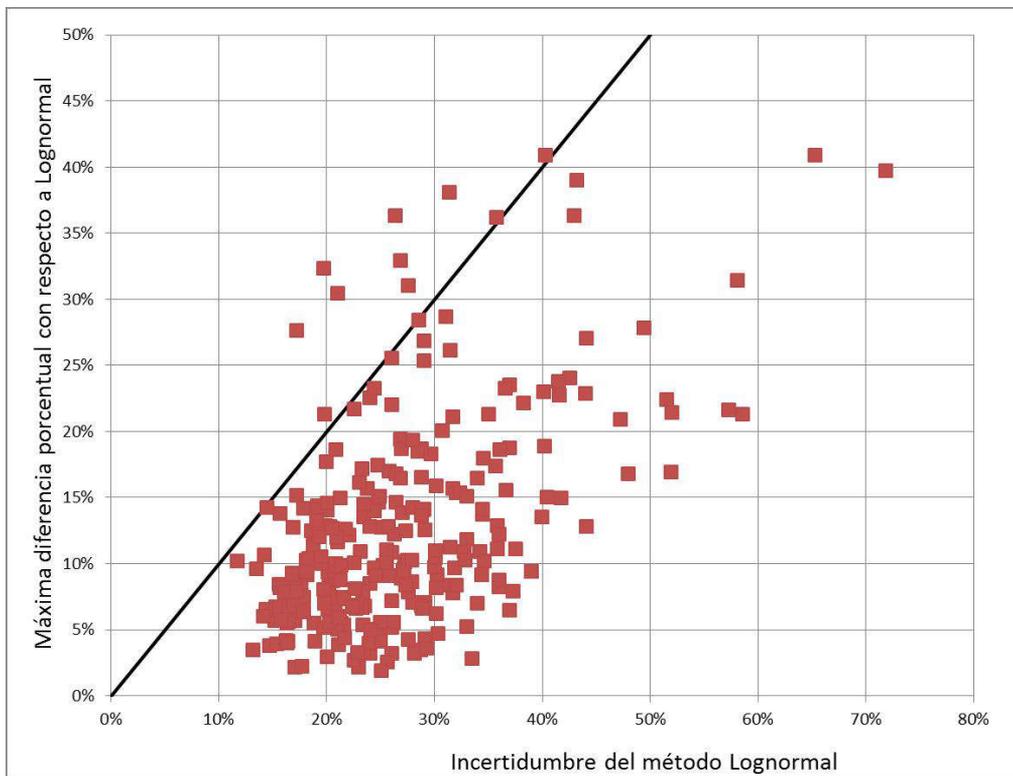


Figura 5. Contraste entre de la máxima diferencia porcentual observada entre los valores estimados con la distribución Log Normal (denominada máxima diferencia porcentual) y la «Incertidumbre del método Lognormal» representativo del tamaño de los intervalos de confianza calculado como la diferencia relativa (límite superior – valor medio) / valor medio.

8. Interpolación de la información puntual utilizando SIG

Con los valores de lámina de lluvia estimados para los distintos periodos de retorno en cada una de las estaciones pluviométricas se utilizaron técnicas de análisis espacial para poder interpolar la información puntual con el objeto de generar grillas de información pluviométrica.

En este proceso se utilizó un Sistema de Información Geográfica (SIG), el cual es definido como un sistema asistido por computadora para la adquisición, almacenamiento, procesamiento, análisis y visualización de datos geográficos. En la actualidad se encuentran disponibles una gran variedad de herramientas y paquetes computacionales para asistir estas tareas. El paquete computacional elegido fue el IdrisiSelva versión 17.00, el cual permite generar un archivo vectorial o de tipo raster georreferenciado. Utilizando esta herramienta se realizó la interpolación espacial de los valores de lámina de lluvia estimados para los distintos periodos de retorno a través del método de Kriging. Específicamente este método permite generar archivos raster a partir de la totalidad de los archivos vectoriales anteriormente creados con la información de lluvia máxima diaria estimada para cada localización. La Figura 6 muestra el Raster interpolado por el método Kriging para la precipitación diaria media máxima anual estimada para un período de retorno de dos años [en mm].

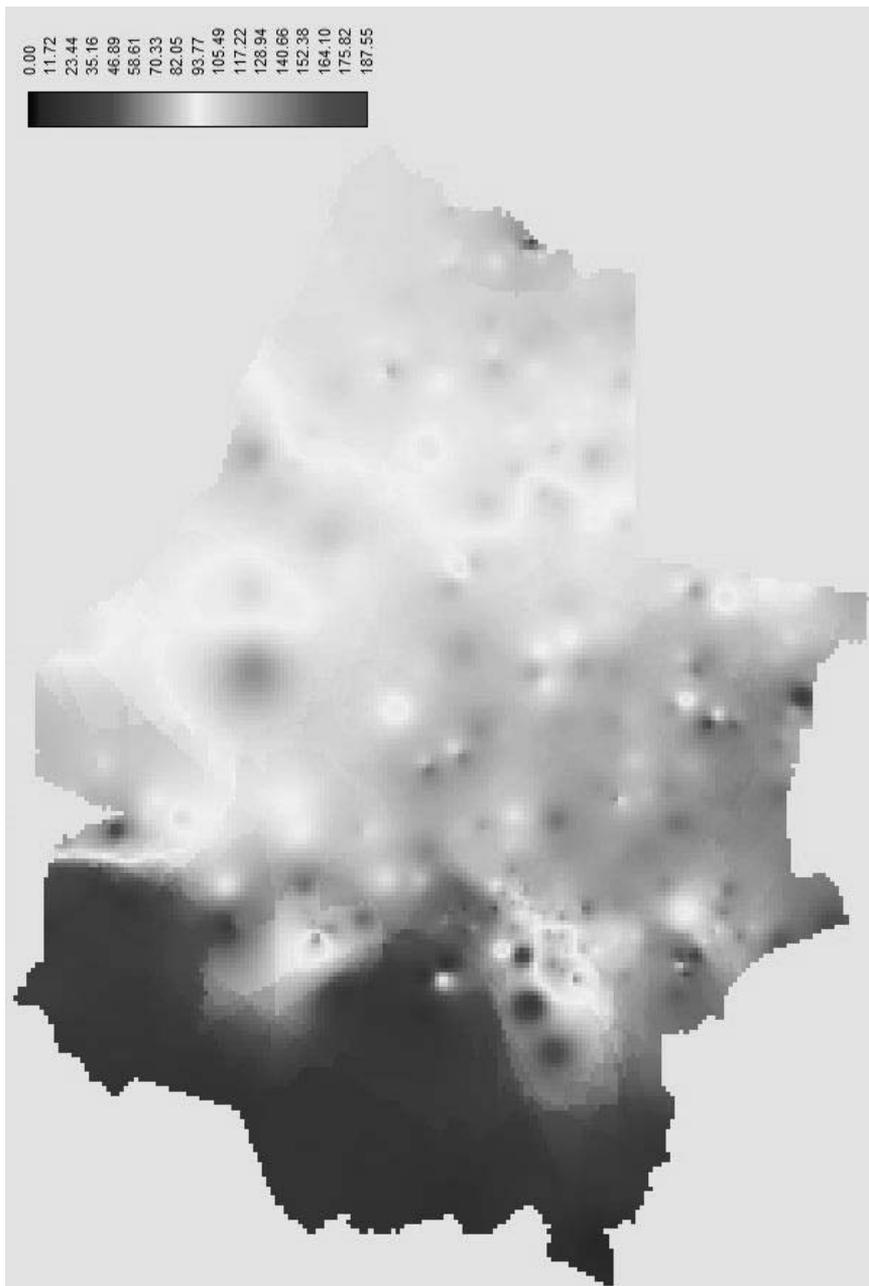


Figura 6. Raster interpolado por el método Kriging para la Precipitación media máxima estimada para un período de retorno de dos años [en mm].

Resultados

En las Figuras 7 a 13 se muestran los mapas de lluvia diaria máxima anual asociados a los diferentes períodos de retorno analizados.

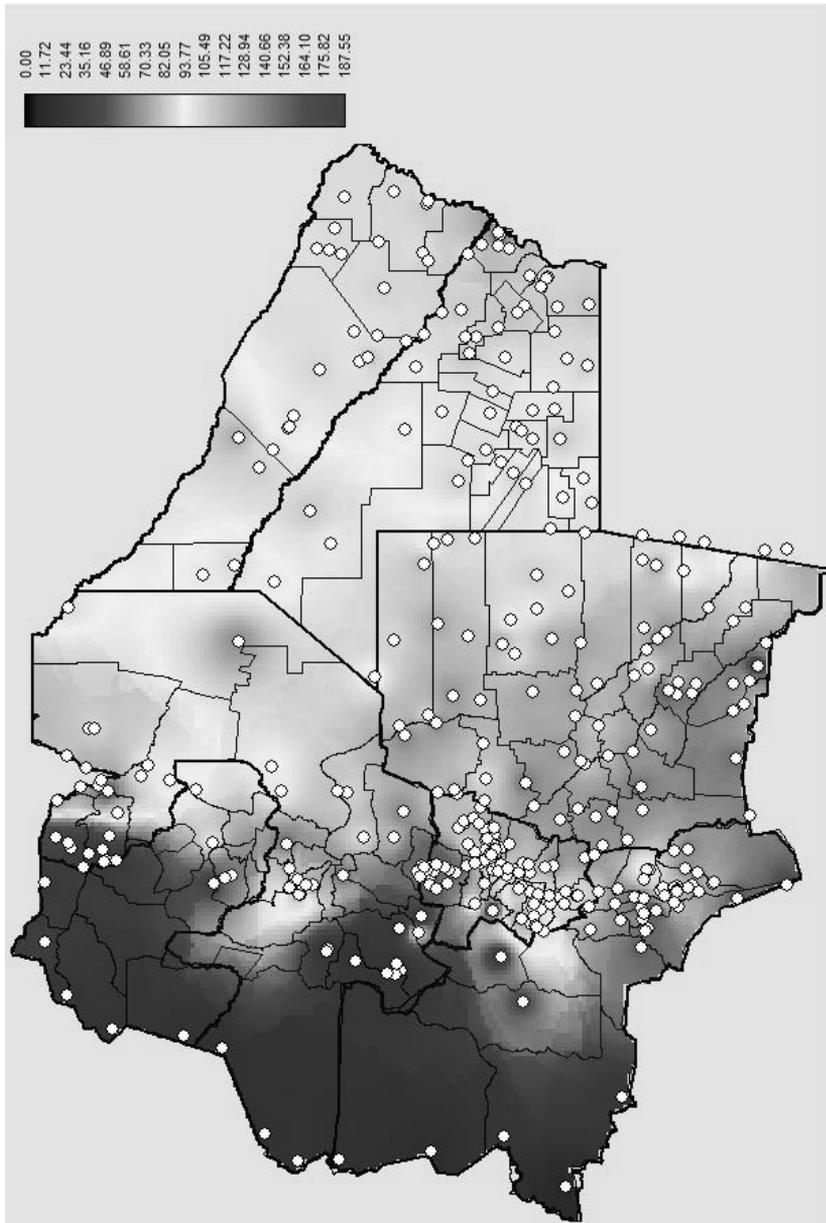


Figura 7. Mapa de lluvias máximas anuales para Período de retorno $T = 2$ años [en mm].

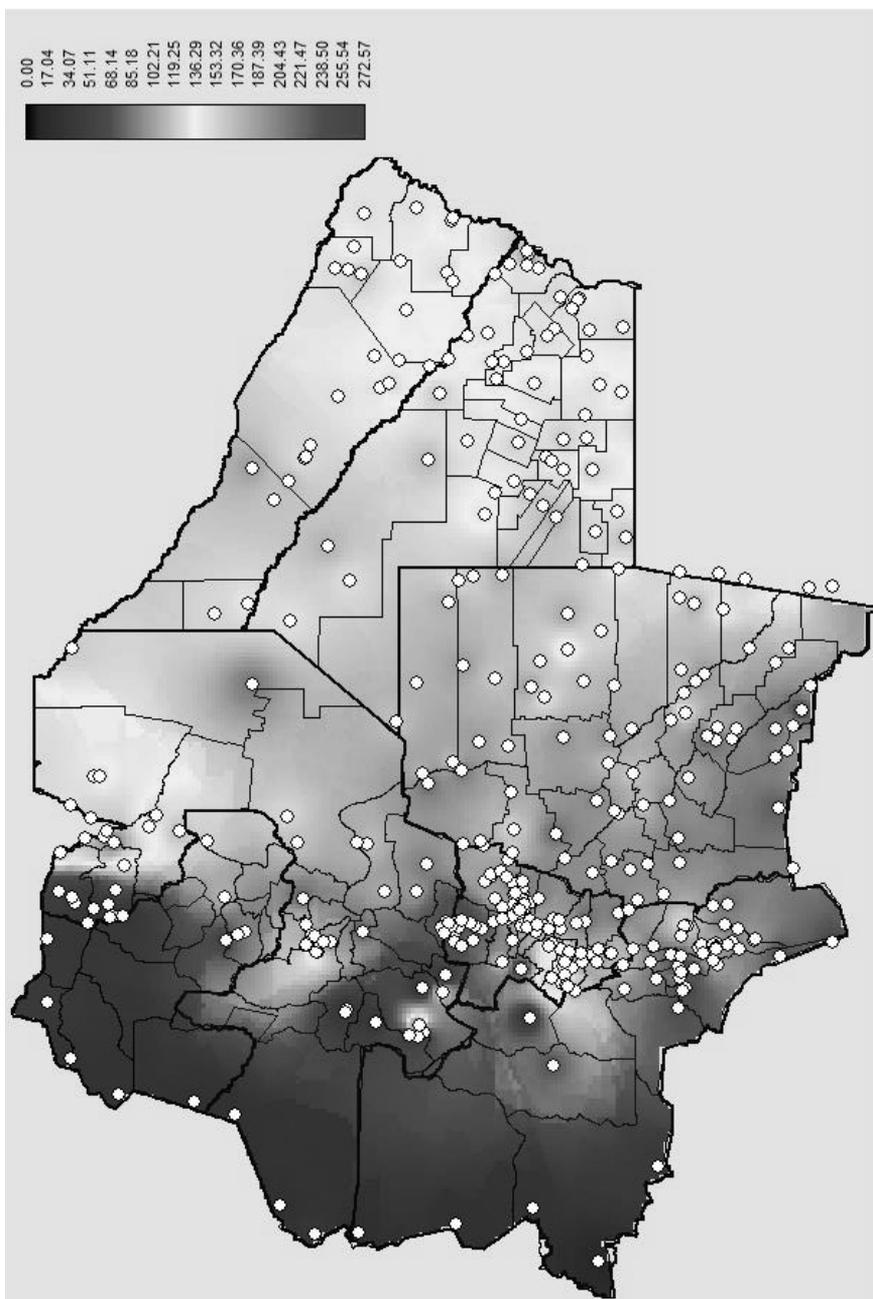


Figura 8. Mapa de lluvias máximas anuales para Periodo de retorno $T = 5$ años [en mm].

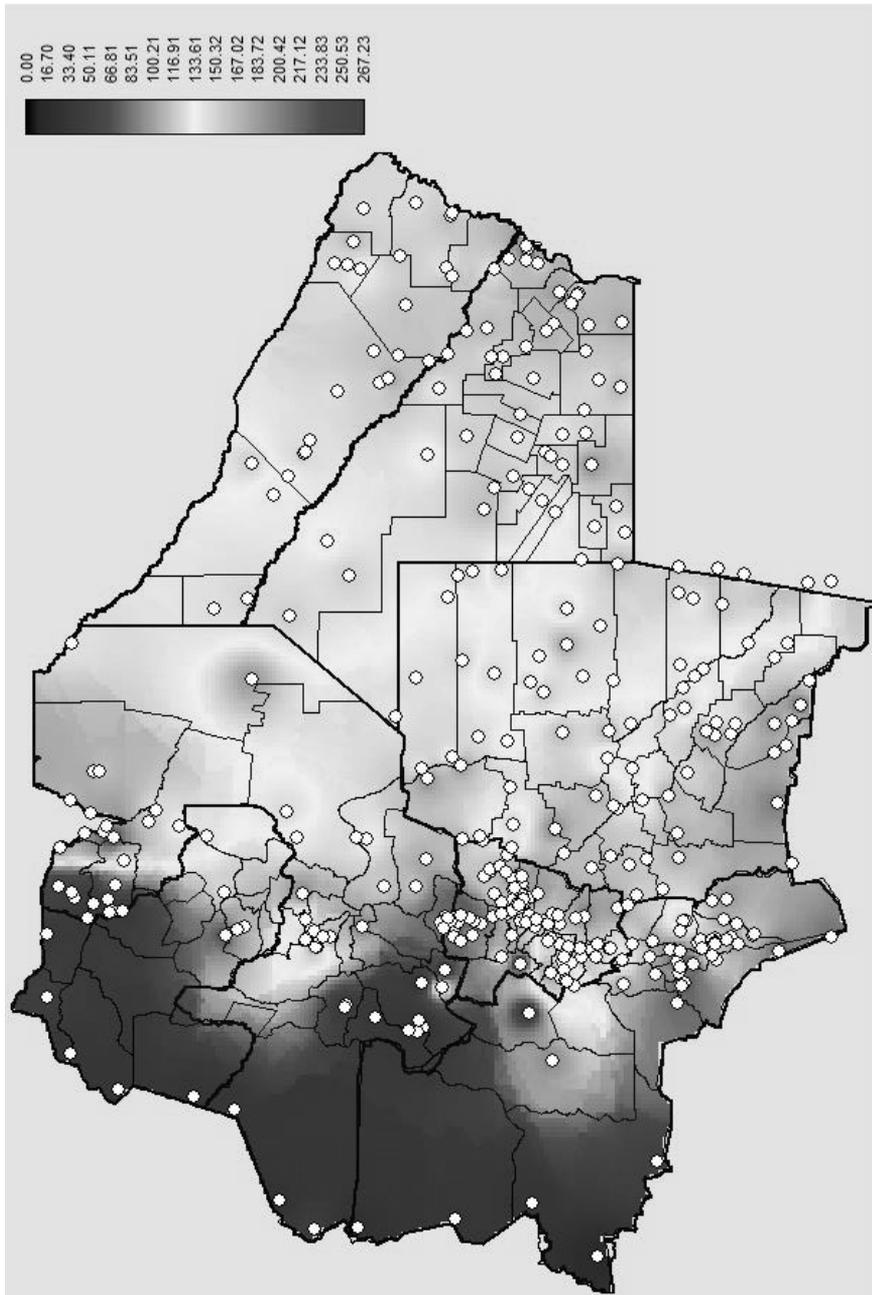


Figura 9. Mapa de lluvias máximas anuales para Periodo de retorno $T = 10$ años [en mm].

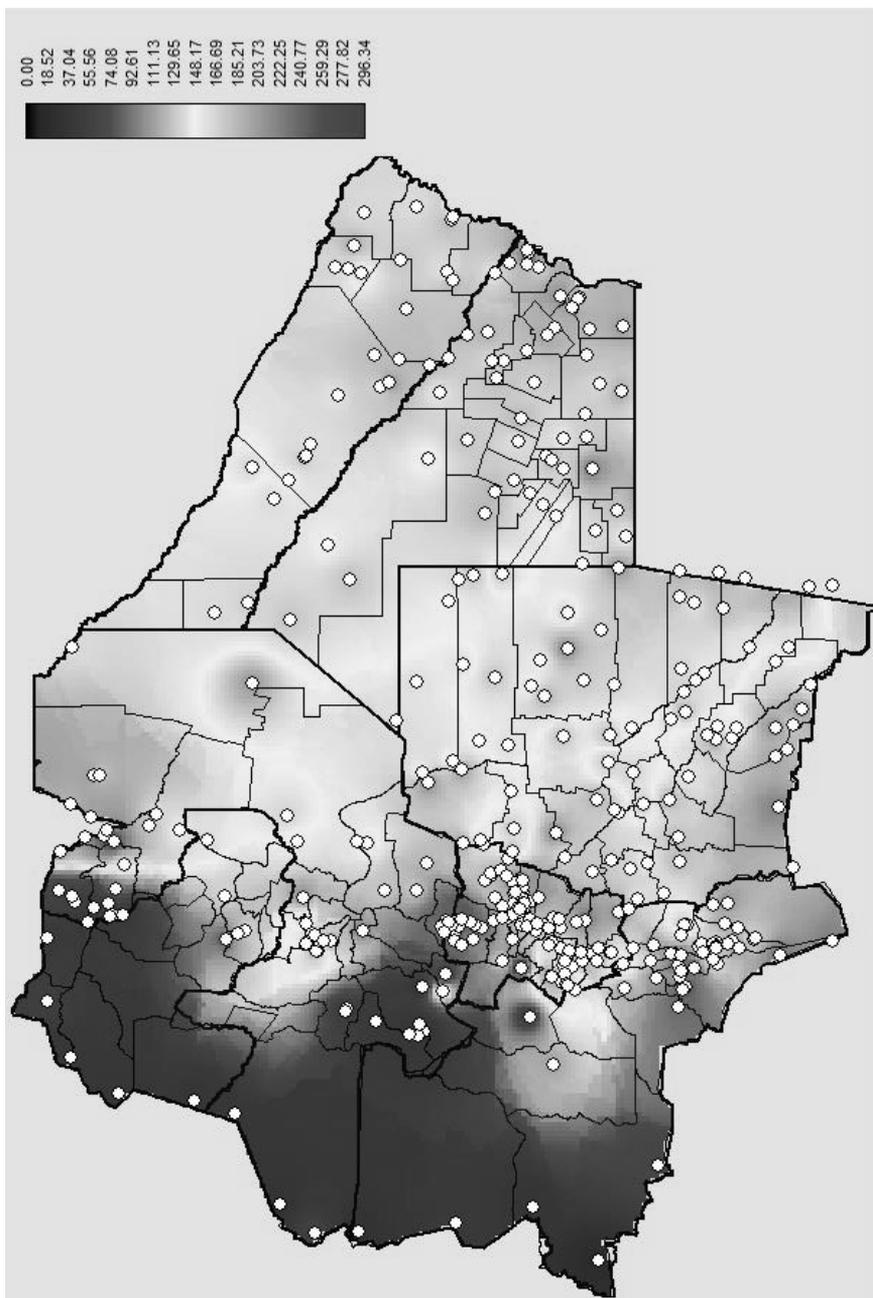


Figura 10. Mapa de lluvias máximas anuales para Periodo de retorno $T = 20$ años [en mm].

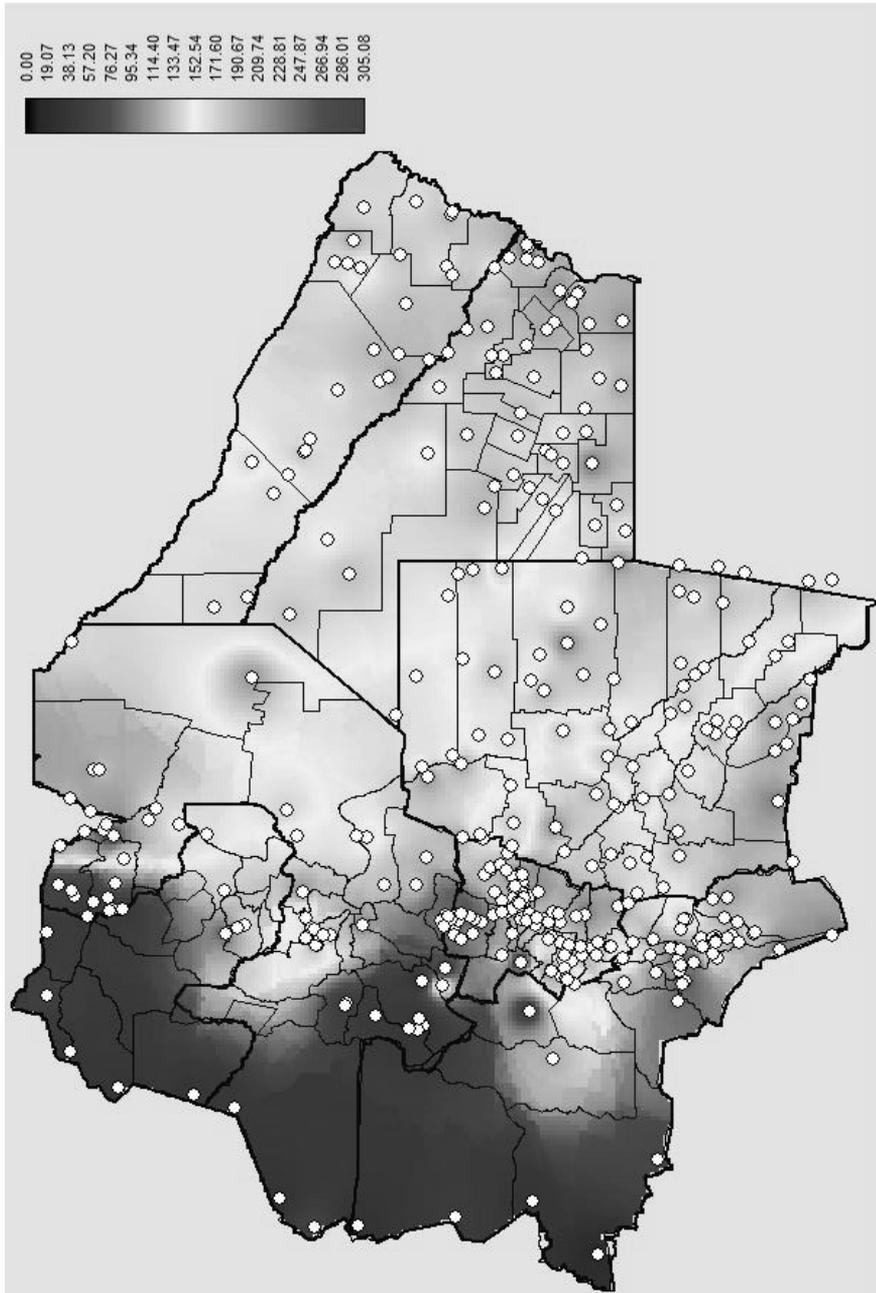


Figura 11. Mapa de lluvias máximas anuales para Período de retorno $T = 25$ años [en mm].

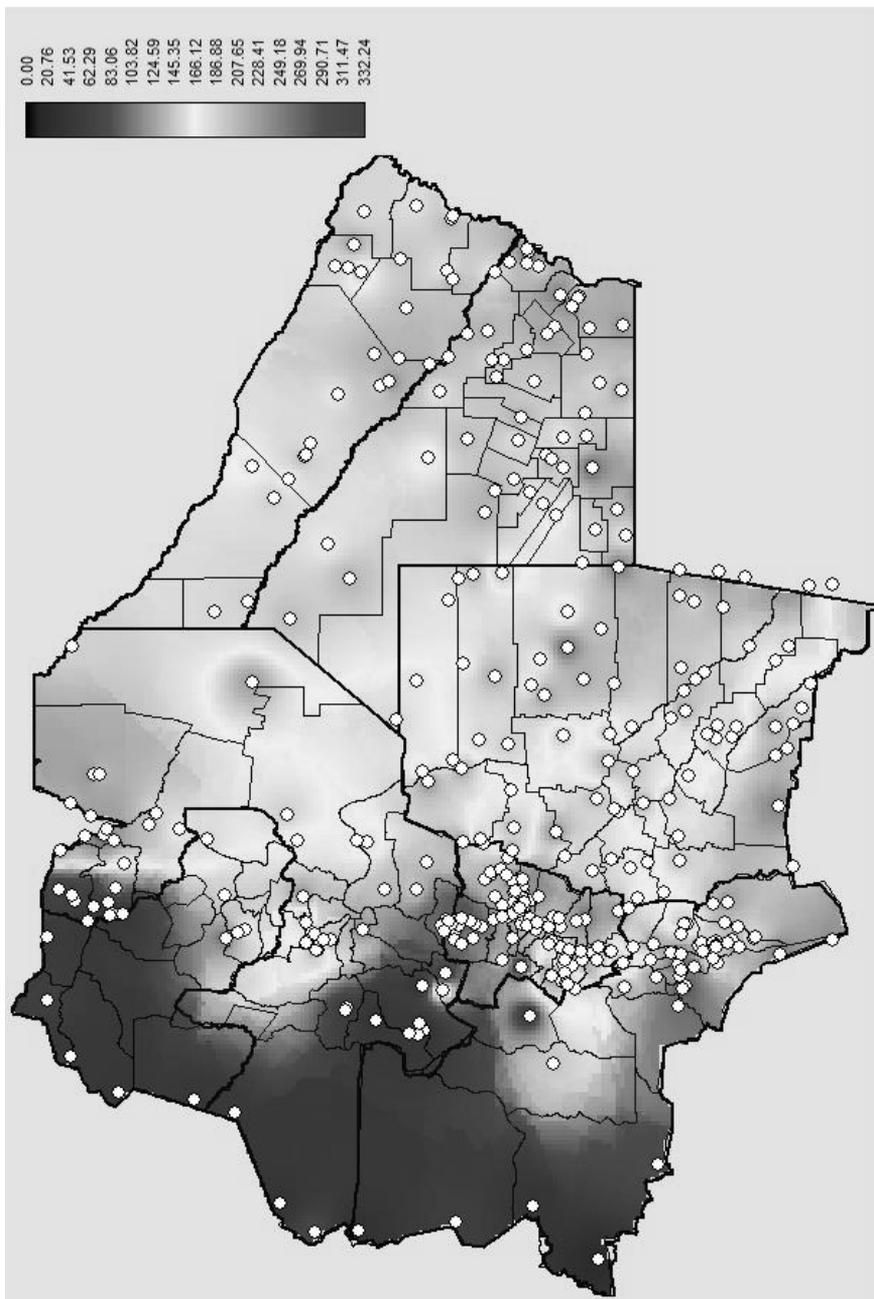


Figura 12. Mapa de lluvias máximas anuales para Periodo de retorno $T = 50$ años [en mm].

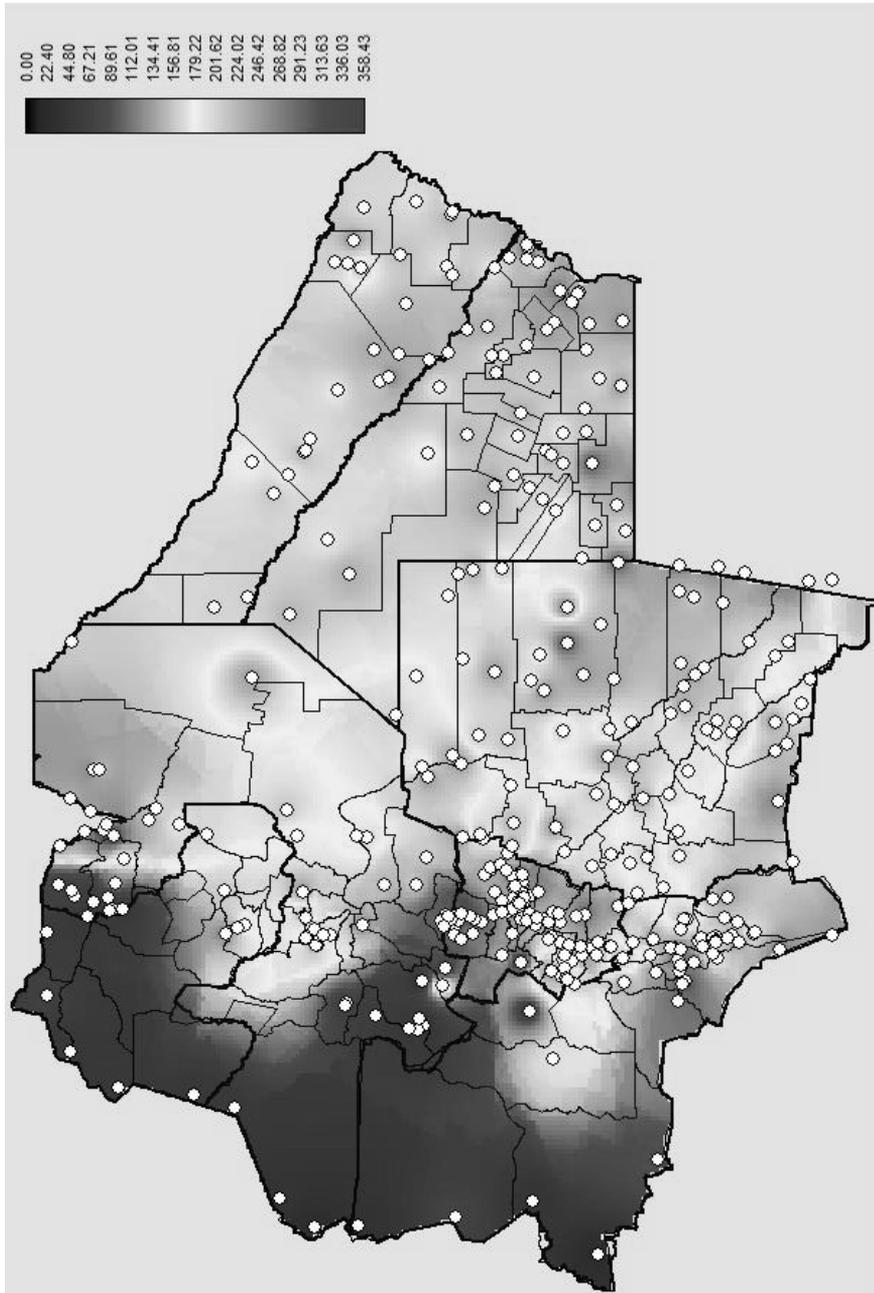


Figura 13. Mapa de lluvias máximas anuales para Periodo de retorno $T = 100$ años [en mm].

Al visualizar estas últimas figuras se observa claramente cómo va creciendo la lámina de lluvia máxima diaria de oeste a este.

Es importante recalcar la importancia de los mapas de lluvias máximas diarias asociadas a diferentes períodos de retorno en toda la zona de estudio construidos en este trabajo. Estos constituyen una herramienta fundamental para la realización de proyectos futuros de dimensionamiento de pequeñas y medianas obras hidráulicas.

Conclusiones

1. Se recopilaron y procesaron registros de 452 estaciones pluviométricas en la región de estudio.
2. Se ubicaron espacialmente las estaciones.
3. La determinación del año hidrológico julio-junio resultó ser el más adecuado para la región de estudio.
4. Se realizó una verificación exhaustiva de las hipótesis estadísticas para todas las estaciones encontradas, dando como resultado que los registros de 278 estaciones verificaron estas pruebas
5. De las 6 alternativas de función de distribución de probabilidad y de método de ajuste ensayados, se adoptó la función LOGNORMAL con parámetros estimados por el método de máxima verosimilitud como representativo y se verificó que la incertidumbre debido al tamaño de las muestras utilizadas es más significativo que la incertidumbre debido al modelo probabilístico utilizado.
6. Se utilizó un Sistema de información Geográfico (SIG) para representar los resultados obtenidos. El programa utilizado fue IDRISI Selva17.00.
7. Se realizó la interpolación espacial a través del método de Kriging, obteniéndose mapas de lluvias máximas diarias asociadas a diferentes períodos de retorno en toda la zona de estudio.

Referencias

- Audagna, Yael (2014). «Regionalización de lluvias máximas diarias en la provincia de Santiago del Estero» - Trabajo final de la carrera de Ingeniería Civil – Facultad de Ciencias Exactas Físicas y Naturales – UNC – Córdoba, Argentina.
- Belli, Virginia Soledad; Cueto, César Agustín; Lovera, Daniela Alejandra; Miguez Ruiz, Martín; Pinasco, Germán Horacio (2012). «Regionalización de lluvias máximas – Aplicación a las provincias de Córdoba y Santa Fe» - Trabajo final de la carrera Ingeniería Civil – Facultad de Ingeniería – Universidad Católica de Córdoba – Córdoba, Argentina.
- Caamaño Nelli, Gabriel; Dasso, Clarita (2003). «Lluvias de Diseño – Conceptos, Técnicas y Experiencias».
- Calderoni, Nahuel (2014). «Regionalización de lluvias máximas diarias en la provincias de Chaco y Formosa» - Trabajo final de la carrera de Ingeniería Civil – Facultad de Ciencias Exactas Físicas y Naturales – UNC – Córdoba, Argentina.

- Catalini, Carlos G.; Maidah, Andrea; García, Carlos M.; Caamaño Nelli, Gabriel (2010). «Mapas digitales de isohietas de lluvias máximas diarias para la provincia de Córdoba».
- Catalini, Carlos G.; Caamaño Nelli, Gabriel E.; García Carlos M., Belli, Virginia Soledad; Bianchiotti, Bernabé; Cueto, César Agustín; Lérída, Gastón; Lovera, Daniela Alejandra; Miguez Ruiz, Martín; Pinasco, Germán Horacio; Sánchez, Ezequiel(2012). «Regionalización Paramétrica de Funciones I-D-T. Aplicación a las provincias de Córdoba y Santa Fe».
- Cervio, Magalí (2014). «Regionalización de lluvias máximas diarias en la provincia de Catamarca» - Trabajo final de la carrera de Ingeniería Civil - Facultad de Ciencias Exactas Físicas y Naturales - UNC - Córdoba, Argentina.
- Guillén, Nicolás Federico (2014). «Estudios Avanzados para el Diseño Hidrológico e Hidráulico de Infraestructura Hídrica» - Tesis de Maestría - Maestría en Ciencias de la Ingeniería Mención Recursos Hídricos - FCEFyN - UNC - Córdoba, Argentina.
- Lencinas, Matías Javier (2014). «Validación de la tormenta de diseño para la Presa Los Molinos - Jujuy - utilizando técnicas de regionalización» - Trabajo final de la carrera de Ingeniería Civil – Facultad de Ciencias Exactas Físicas y Naturales - UNC - Córdoba, Argentina.

Analysis of Gutenberg-Richter b -value and m_{\max} Part I: Exact Solution of Kijko-Sellevoll Estimator of m_{\max}

Mika Haarala^{1,2} y Lía Orosco^{1,3}

Abstract

This report is the first of a series of three which have the main goal to achieve a method to estimate the parameters β and m_{\max} that are essential when Gutenberg-Richter law is used for seismic hazard assessment.

We give the exact solution of Kijko-Sellevoll approach to estimate and also a new method to calculate m_{\max} applying series. It proved to be numerically more stable even in the cases when wide range of magnitudes or large catalogue is used.

Keywords: Mmax - b -value - Gutenberg-Richter distribution function - Kijko-Sellevoll estimator - series

Resumen

Este es el primero de una serie de tres informes sobre un trabajo que tiene como principal objetivo lograr un método para estimar los parámetros β y m_{\max} , que son esenciales cuando se utiliza la ley de Gutenberg – Richter para la estimación de la peligrosidad sísmica.

Proponemos la solución exacta del método de Kijko-Sellevoll para estimar m_{\max} como así también mostramos un nuevo método para calcular m_{\max} aplicando series. Este método es numéricamente más estable, aun en los casos en que se utiliza un catálogo sísmico en el que los valores se ubican en un intervalo amplio de magnitudes.

Palabras clave: Mmax - b - función de distribución Gutenberg-Richter - estimador de Kijko-Sellevoll - series

Introduction

In seismic hazard assessment studies, the very well-known frequency-magnitude distribution (Ishimoto, Iida, 1939; Gutenberg, Richter, 1944), commonly known as Gutenberg-Richter law,

$$\log_{10} N(M) = a - bM$$

¹ Instituto de Estudios Interdisciplinarios de Ingeniería (IESIING) – Facultad de Ingeniería – UCASAL

² Acrenet Oy, Joensuu, Finland.

³ Facultad de Ingeniería. Universidad Nacional de Salta

is the cumulative number of events with magnitude greater or equal than M , a and b are some unknown constants to be determined by some method. When a probabilistic approach is used, the Gutenberg – Richter probabilistic density function

$$f(m) = \beta \exp[-\beta(m - m_{\min})] \quad (1)$$

or the double truncated Gutenberg-Richter distribution function

$$f(m) = \frac{\beta \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \quad (2)$$

are still applied and investigated (Anagnostopoulos et al., 2008; Ishibe, Shimazaki, 2008; Kahraman et al., 2008; Leyton et al., 2009; Amorèse et al., 2010; Holschneider et al., 2011; Zúñiga, Figueroa-Soto, 2012; Kijko, Smit, 2012; Mostafanejad et al., 2013; Rong et al., 2014; Márquez-Ramírez et al., 2015).

In (1) and (2) the β -value is related with b as $\beta = b \log(10)$, m_{\min} is known like threshold of completeness of seismic catalogue and m_{\max} is the maximum earthquake probably to occur.

In 1965, using moments of Gamma distribution function, Utsu (1965) derived a simple estimator for β -value, in the case of unbounded expression (1). Aki (1965) showed that this estimator is also a maximum likelihood estimator for the Gutenberg-Richter distribution function. In this work we will call it as Aki-Utsu estimator, which has been quite popular because of its simplicity.

Page (1968) proposed a maximum likelihood estimator for the bounded expression (2) which needs to be solved iteratively and it needs to estimate somehow the parameter m_{\max} . Normally this is set to maximum observed magnitude.

Hamilton (1967), studying the stability of mean value and variance of sequences of earthquakes, used the moment method moments; later, Cosentino (Cosentino, Luzio, 1976; Cosentino et al., 1977) published the moment method estimators for the Gutenberg-Richter distribution function.

In 1984, Kijko (1984) presented for the first time his idea to calculate the estimator of m_{\max} . Later, Kijko and Sellevoll (1989) developed the method itself using the double truncated Gutenberg-Richter distribution function (2). Kijko and Graham (1998), with basis in Cramer's approximation (Cramer, 1961), generalized this method making possible to apply it for different distribution functions. Kijko (2004) called the estimator of m_{\max} of the Gutenberg-Richter model (2) as a «Kijko-Sellevoll (KS) estimator».

When applying Gutenberg-Richter model, parameter β must be defined a priori with some hypothesis; m_{\max} is set as the maximum observed magnitude, as infinity or that determined by empirical formula when possible.

The objectives of this work are to show the algebraic solution to KS estimator, and to propose a method to estimate m_{\max} . This work is the first part of a series of three, which have the main goal to achieve a method to estimate β and m_{\max} by using exact solution of Gutenberg-Richter law.

Exact solution of Kijko-Sellevoll estimator

Firstly we give the exact solution of KS model using exact value of β . Actually K-S estimator has two different solutions: first one (Kijko-Sellevoll function 1, named as KS-1) is a solution of the original KS model and the second one (Kijko-Sellevoll function 2, named as KS-2) is its counterpart such as $f_n^{KS-1} + f_n^{KS-2} = \beta(m_{max} - m_{min})$

In this work we are analyzing the double truncated Gutenberg-Richter distribution function (2) which has cumulative distribution function (CDF)

$$F_M(m | m_{max}) = \begin{cases} 0, & \text{for } m < m_{min}, \\ \frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} & \text{for } m_{min} \leq m < m_{max}, \\ 1, & \text{for } m_{max} \leq m. \end{cases} \quad (3)$$

Because the unbounded limit of this distribution function exists

$$F_M(m | m_{max} = \infty) = \lim_{m_{max} \rightarrow \infty} F_M(m | m_{max}) = \begin{cases} 0, & \text{for } m < m_{min}, \\ 1 - \exp[-\beta(m - m_{min})] & \text{for } m_{min} \leq m, \end{cases}$$

and the lower limit exists

$$F_M(m | m_{max} = m_{min}) = \lim_{m_{max} \rightarrow m_{min}} F_M(m | m_{max}) = \begin{cases} 0, & \text{for } m < m_{min}, \\ 1, & \text{for } m_{min} \leq m, \end{cases}$$

we assume that $m_{max} \in [m_{min}, \infty]$

Let $M_1, M_2, \dots, M_n \in [m_{min}, m_{max}]$ be a set of random variables (which we shall call catalogue C_n of size n). Let $M_{(1)} \leq M_{(2)} \leq \dots \leq M_{(n)}$ denote the ordered values of M_1, M_2, \dots, M_n . That is to say, the random variable $M_{(n)}$ is a maximum in the catalogue C_n . We assume also that these random variables are independently and identically distributed (IID) with CDF $F_M(m)$ given by (3). Let $m_{(1)} \leq m_{(2)} \leq \dots \leq m_{(n)}$ be an ordered sample of magnitudes where $m_{(1)}$ is a minimum observed magnitude ($m_{min} \leq m_{(1)}$) and $m_{(n)}$ is a maximum observed magnitude ($m_{(n)} \leq m_{max}$). This $m_{(n)}$ has a CDF

$$F_{M_{(n)}}(m | m_{max}) = \begin{cases} 0, & \text{for } m < m_{min}, \\ [F_M(m | m_{max})]^n & \text{for } m_{min} \leq m < m_{max}, \\ 1, & \text{for } m_{max} \leq m. \end{cases} \quad (4)$$

Integrating by parts, the expected value of $M_{(n)}$ is

$$E\left(M_{(n)} \mid m_{\max}\right) = \int_{m_{\min}}^{m_{\max}} m dF_{M_{(n)}}(m \mid m_{\max}) = m_{\max} - \int_{m_{\min}}^{m_{\max}} F_{M_{(n)}}(m \mid m_{\max}) dm. \quad (5)$$

Then Kijko set

$$m_{\max} = E\left(M_{(n)} \mid m_{\max}\right) + \int_{m_{\min}}^{m_{\max}} F_{M_{(n)}}(m \mid m_{\max}) dm.$$

(In Appendix A we show the Kijko's method to find the estimator \hat{m}_{\max} .)

We define a new function

$$g(\mathfrak{M}) = \left[\mathfrak{M} - E\left(M_{(n)} \mid m_{\max}\right) \right] - \int_{m_{\min}}^{\mathfrak{M}} F_{M_{(n)}}(m \mid \mathfrak{M}) dm, \quad (6)$$

where $\mathfrak{M} \in [m_{\min}, \infty]$. For the estimator \hat{m}_{\max} holds $g(\hat{m}_{\max}) = 0$ because it is a solution of equation (5). Function g is negative at the point $\mathfrak{M} = E\left(M_{(n)} \mid m_{\max}\right)$ since in equation (6) the first term is zero and, because of $F_{M_{(n)}}(m \mid \mathfrak{M})$ is a positive function for any fixed n (i.e. $n < \infty$), the integral is positive.

The derivative of function g in (6) is (using Leibniz's theorem for differentiation of an integral; Abramowitz, Stegun, 1972)

$$\begin{aligned} g'(\mathfrak{M}) &= 1 - F_{M_{(n)}}(\mathfrak{M} \mid \mathfrak{M}) - \int_{m_{\min}}^{\mathfrak{M}} \frac{\partial}{\partial \mathfrak{M}} F_{M_{(n)}}(m \mid \mathfrak{M}) dm \\ &= -n \int_{m_{\min}}^{\mathfrak{M}} \left[F_M(m \mid \mathfrak{M}) \right]^{n-1} \frac{\partial}{\partial \mathfrak{M}} F_M(m \mid \mathfrak{M}) dm \geq 0 \end{aligned} \quad (7)$$

for all $\mathfrak{M} \in]m_{\min}, \infty[$ since $\frac{\partial}{\partial \mathfrak{M}} F_M(m \mid \mathfrak{M}) \leq 0$. Thus g is a monotonically increasing function and because we showed above that there exists at least one point where the function g is negative, so it has at most one solution for $g(\hat{m}_{\max}) = 0$. This makes attractive the idea to apply the Newton-Raphson method to find the estimator \hat{m}_{\max} .

To find the solution of the last integral in equation (5) we write

$$\Delta = \int_{m_{\min}}^{m_{\max}} F_{M_{(n)}}(m \mid m_{\max}) dm = \frac{\int_{m_{\min}}^{m_{\max}} \left(1 - \exp[-\beta(m - m_{\min})]\right)^n dm}{\left(1 - \exp[-\beta(m_{\max} - m_{\min})]\right)^n}. \quad (8)$$

We can calculate the derivative for $\left(1 - \exp[-\beta(m - m_{\min})]\right)^n$ as

$$-\frac{1}{\beta n} \frac{\partial}{\partial m} (1 - \exp[-\beta(m - m_{min})])^n = (1 - \exp[-\beta(m - m_{min})])^n - (1 - \exp[-\beta(m - m_{min})])^{n-1},$$

which gives the integration formula

$$\int_{m_{min}}^{m_{max}} (1 - \exp[-\beta(m - m_{min})])^n dm = -\frac{1}{\beta n} (1 - \exp[-\beta(m_{max} - m_{min})])^n + \int_{m_{min}}^{m_{max}} (1 - \exp[-\beta(m - m_{min})])^{n-1} dm.$$

Applying this n times we can eliminate the power (on the last step the power is equal to zero) and the integral (8) can be expressed by

$$\Delta = \frac{1}{\beta} \frac{\beta(m_{max} - m_{min}) - \sum_{k=1}^n \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k}}{(1 - \exp[-\beta(m_{max} - m_{min})])^n}. \tag{9}$$

To the logarithm function holds (Abramowitz, Stegun, 1972)

$$-\log(1 - z) = \sum_{k=1}^{\infty} \frac{z^k}{k}, \quad |z| \leq 1, z \neq 1. \tag{10}$$

We can write now

$$\beta(m_{max} - m_{min}) = -\log[1 - (1 - \exp[-\beta(m_{max} - m_{min})])].$$

Setting $z = 1 - \exp[-\beta(m_{max} - m_{min})]$ we see that conditions of (10) i.e. $0 \leq z < 1$ holds when $m_{max} \in [m_{min}, \infty[$ and applying this to equation (10) we have

$$\beta(m_{max} - m_{min}) = \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k}. \tag{11}$$

Hence the sum in (9) represents the first n terms of the series (11) so we can rewrite(9) in the form

$$\Delta = \frac{1}{\beta} \frac{\sum_{k=n+1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k}}{(1 - \exp[-\beta(m_{max} - m_{min})])^n}$$

Dividing each term by denominator and re-indexing the series we get the final result

$$\Delta = \frac{1}{\beta} \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^k}{k + n} \tag{12}$$

This is equal to the solution (9) when n is an integer. In the series (12) the variable n is continuous ($n \in \mathbb{R}$) and it has a derivative, whereas the solution (9) has not it. In Appendix B we give the method to calculate directly the numerical values for the series (12). In normal cases when n is small and $m_{\max} - m_{\min}$ is enough big, we can get numerical results of the series by using the solution (9). Any way it is recommended to use the method of series given in appendix B to avoid the numerical instability of the formula (9).

Now it is possible to rewrite the expression of expected value (5) as

$$\beta [m_{\max} - E(M_{(n)} | m_{\max})] = \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^k}{k + n} \tag{13}$$

The right side is the Kijko-Sellevoll function 1 (KS-1 or f_n^{KS-1}). In fact the actual form of the function is $y = f_n^{KS-1}(x)$ where $x = \beta(m_{\max} - m_{\min})$ and $y = \beta [m_{\max} - E(M_{(n)} | m_{\max})]$. Explanation of this is abstract. If we change the variable in the distribution function (3) setting $m = x/\beta + m_{\min}$, with $\beta > 0$ we get a normalized CDF

$$\tilde{F}_X(x | x_{\max}) = \begin{cases} 0, & \text{for } x < 0, \\ \frac{1 - \exp(-x)}{1 - \exp(-x_{\max})} & \text{for } 0 \leq x < x_{\max}, \\ 1, & \text{for } x_{\max} \leq x. \end{cases}$$

where $x_{\max} = \beta(m_{\max} - m_{\min})$ is a pseudo maximum magnitude. This is the CDF of truncated exponential distribution function and the KS-1 function works in this space. It is to say that the KS function measures the relation between the pseudo maximum and the pseudo expected value of maximum.

We pointed out before that there is another KS function. Our goal was to establish if the derivative of g in the equation (6) can be written by using same series than in the case of KS-1. Fortunately we found out that it is possible to write g and g' using same series, for example by means of the use of Kummer's transformation (see Abramowitz and Stegun, 1972).

We start with the counterpart of the expected value given by (5), which we will write as

$$E(M_{(n)} | m_{\max}) = m_{\min} + \int_{m_{\min}}^{m_{\max}} (1 - F_{M_{(n)}}(m | m_{\max})) dm \tag{14}$$

If the relation (5) measures the probability of occurrence, equation (14) measures the probability

of non-occurrence. Because of $1 - F_{M_{(n)}}(m | m_{max})$ is positive function when $m_{min} < m_{max}$, then $E(M_{(n)} | m_{max}) > m_{min}$ for all fixed n and it is zero only if $m_{min} = m_{max}$.

Solving the integral we get (14)

$$E(M_{(n)} | m_{max}) = m_{min} + (m_{max} - m_{min}) - \int_{m_{min}}^{m_{max}} F_{M_{(n)}}(m | m_{max}) dm.$$

If we eliminate the minimum m_{min} we shall get back the formula (5), but we replace the difference $m_{max} - m_{min}$ with the series (11) and integral with the series (12) instead, so we get

$$E(M_{(n)} | m_{max}) = m_{min} + \frac{1}{\beta} \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k} - \frac{1}{\beta} \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k+n}.$$

To each term it holds

$$\frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k} - \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k+n} = n \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k(k+n)}$$

so it implies that

$$\beta [E(M_{(n)} | m_{max}) - m_{min}] = n \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k(k+n)}. \tag{15}$$

Here the right side is the Kijko-Sellevoll function 2 (KS-2 or f_n^{KS-2}). The relation between KS-1 and KS-2 is

$$\begin{aligned} \beta(m_{max} - m_{min}) &= \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k+n} + n \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k(k+n)} \\ &= f_n^{KS-1}(\beta(m_{max} - m_{min})) + f_n^{KS-2}(\beta(m_{max} - m_{min})). \end{aligned} \tag{16}$$

Using this relationship, we can find the finite sum formula for KS-2

$$f_n^{KS-2}(\beta(m_{max} - m_{min})) = \frac{\left[(1 - \exp[-\beta(m_{max} - m_{min})])^n - 1 \right] \beta(m_{max} - m_{min}) + \sum_{k=1}^n \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{k}}{(1 - \exp[-\beta(m_{max} - m_{min})])^n} \tag{17}$$

The expression (17) can be used to calculate numerical values to the function KS-2. Any way as we said before (in the case KS-1), it is recommended to use the numerical solution of series instead of formula (17) because of numerical stability. We could also calculate KS-2 when we know KS-1 by means of relation (16).

The function KS-2 gives more «natural way» the solution of m_{\max} the than the function KS-1 because there is some inverse function such that

$$f^{-1}\left(\beta\left[E\left(M_{(n)}\mid m_{\max}\right)-m_{\min}\right]\right)=\beta\left(m_{\max}-m_{\min}\right).$$

Unfortunately, we cannot say what is the inverse function.

When maximum m_{\max} is considered to be infinity, and taking into account KS-1 function for fixed $n \geq 0$, we have

$$\lim_{m_{\max} \rightarrow \infty} \sum_{k=1}^{\infty} \frac{\left(1 - \exp\left[-\beta\left(m_{\max} - m_{\min}\right)\right]\right)^k}{k+n} = \sum_{k=1}^{\infty} \frac{1}{k+n} = \sum_{k=1}^{\infty} \frac{1}{k} - \sum_{k=1}^n \frac{1}{k} \rightarrow \infty,$$

so the KS-1 diverges because the harmonic series diverges. Hence the KS-1 is an unbounded function.

The case of the KS-2 function is different. We need some results of Psi function (Abramowitz and Stegun, 1972)

$$\psi(1+z) = -\gamma + \sum_{k=1}^{\infty} \frac{z}{k(k+z)}, \quad z \neq -1, -2, -3, \dots,$$

$$\psi(1) = -\gamma, \quad \psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}, \quad n \geq 2,$$

where γ is the Euler-Mascheroni constant. When n is zero, the KS-2 function is zero. Let's assume that $n \geq 0$ is fixed then the KS-2 function converges since it is positive term series and

$$\lim_{m_{\max} \rightarrow \infty} \sum_{k=1}^{\infty} \frac{n\left(1 - \exp\left[-\beta\left(m_{\max} - m_{\min}\right)\right]\right)^k}{k(k+n)} = n \sum_{k=1}^{\infty} \frac{1}{k(k+n)} \leq n \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty \tag{18}$$

because the last series converges. Moreover, if n is positive integer ($n \geq 1$) then

$$\sum_{k=1}^{\infty} \frac{n}{k(k+n)} = \psi(1+n) + \gamma = \sum_{k=1}^n \frac{1}{k} = H_n, \tag{19}$$

where H_n is called the harmonic number. Thus the KS-2 is a bounded function for any fixed n .

If we apply the results(18) and (19) to the equation we get

$$E\left(M_{(n)}\mid \infty\right)-m_{\min} = \frac{H_n}{\beta}. \tag{20}$$

Because of we assumed that the value β is constant, the right side of (20) is constant for any fixed n . The formula (20) shows a fundamental phenomenon of unbounded distribution function. No matter how we choose the minimum m_{min} , the distance to the expected maximum value is always the same.

Estimator for m_{max}

As we showed there are two different solutions for the estimator $E(M_{(n)} | m_{min})$

$$\begin{aligned} E(M_{(n)} | m_{max}) &= m_{max} - \frac{1}{\beta} f_n^{KS-1}(\beta(m_{max} - m_{min})) \\ &= m_{min} + \frac{1}{\beta} f_n^{KS-2}(\beta(m_{max} - m_{min})). \end{aligned}$$

Also our auxiliary function can (6) be written using the KS-1 or KS-2 function

$$\begin{aligned} g(\mathfrak{M}) &= \mathfrak{M} - E(M_{(n)} | m_{max}) - \frac{1}{\beta} f_n^{KS-1}(\beta(\mathfrak{M} - m_{min})) \\ &= m_{min} - E(M_{(n)} | m_{max}) + \frac{1}{\beta} f_n^{KS-2}(\beta(\mathfrak{M} - m_{min})). \end{aligned}$$

Since they present the same function, they have the same derivative. We get

$$\begin{aligned} g'(\mathfrak{M}) &= \frac{1}{\beta} \frac{\partial}{\partial \mathfrak{M}} f_n^{KS-2}(\beta(\mathfrak{M} - m_{min})) \\ &= n \frac{\exp[-\beta(\mathfrak{M} - m_{min})]}{1 - \exp[-\beta(\mathfrak{M} - m_{min})]} f_n^{KS-1}(\beta(\mathfrak{M} - m_{min})). \end{aligned}$$

Now the step of Newton-Raphson method (NRM) has the expression

$$\begin{aligned} \mathfrak{M}_0 &= E(M_{(n)} | m_{max}), \\ \mathfrak{M}_{k+1} &= \mathfrak{M}_k - \frac{g(\mathfrak{M}_k)}{g'(\mathfrak{M}_k)} = \mathfrak{M}_k - \frac{\mathfrak{M}_k - E(M_{(n)} | m_{max}) - \frac{1}{\beta} f_n^{KS-1}(\beta(\mathfrak{M}_k - m_{min}))}{n \frac{\exp[-\beta(\mathfrak{M}_k - m_{min})]}{1 - \exp[-\beta(\mathfrak{M}_k - m_{min})]} f_n^{KS-1}(\beta(\mathfrak{M}_k - m_{min}))} \\ &= \mathfrak{M}_k + \frac{\exp[\beta(\mathfrak{M}_k - m_{min})] - 1}{n\beta} \left\{ 1 - \frac{\beta[\mathfrak{M}_k - E(M_{(n)} | m_{max})]}{f_n^{KS-1}(\beta(\mathfrak{M}_k - m_{min}))} \right\}. \end{aligned}$$

The second term in the difference between brackets, measures the distance to the exact solution when it is different than one. This step is a Pisarenko estimator (Pisarenko et al., 1996) or a Tate-Pisarenko estimator (21) (Kijko and Graham, 1998), at $\mathfrak{m}_0 = E(M_{(n)} | m_{\max})$ (using the estimator of the maximum observed magnitude $m_{(n)}$)

$$\hat{m}_{\max} = m_{(n)} + \frac{1}{n\beta} \frac{1 - \exp[-\beta(m_{(n)} - m_{\min})]}{\exp[-\beta(m_{(n)} - m_{\min})]} \quad (21)$$

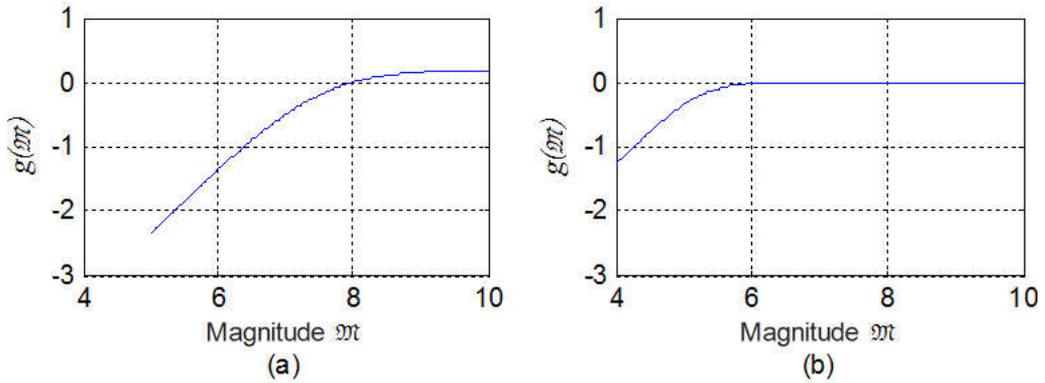


Figure 1: Function with $g(\mathfrak{m})$ with different set of parameters.

We show in figure 1 how the function g (in ordinates) varies with magnitude (in abscissas); it looks like Γ . The figure 1a has been drawn with parameters $b = 1, m_{\max} = 8, m_{\min} = 5$ and $n = 200$, and the figure 1b with $b = 2, m_{\max} = 9.5, m_{\min} = 4$ and $n = 200$. As we can see in case 1b the derivative is almost zero and NRM cannot solve it using double-precision arithmetic. In simulations the NRM can find the solution up to $b(m_{\max} - m_{\min}) = 7$ with exact expected value, but it is numerically stable up to $b(m_{\max} - m_{\min}) = 4.5$.

The artificial catalogue can be generated by using inverse function of (3)

$$m_i = m_{\min} - \frac{1}{\beta} \log \left[1 - (1 - \exp[-\beta(m_{\max} - m_{\min})]) u_i \right], \quad (22)$$

where $u_i \in U([0,1])$ is uniformly distributed random variable between zero and one. From this artificial catalogue we choose the maximum observed value $m_{(n):k} = \max \{m_1, \dots, m_n\}$, $m_i \in C_{n:k}$.

In order to get some estimator of the expected value $E(M_{(n)} | m_{\max})$, we could have

$$\bar{m}_{\max} = m_{\min} + \frac{1}{\beta N} \sum_{k=1}^N (f_n^{KS-2})^{-1} \left(\beta [m_{(n):k} - m_{\min}] \right),$$

where \bar{m}_{\max} is the mean value of the estimators of m_{\max} . In this case we use a single maximum event to estimate the expected value of maximum to each catalogue and finally we calculate the mean of maximums.

On the other hand, we can calculate first the mean value of maximums of all catalogues, such like

$$\bar{m}_{(n)} = \frac{1}{N} \sum_{k=1}^N m_{(n):k} \quad (23)$$

and consider the expression(23) as the estimator of the expected value of the maximum $E(M_{(n)} | m_{\max})$. Thus, we obtain the estimator of m_{\max} using the mean value of maximums of all catalogues

$$\hat{m}_{\max} = m_{\min} + \frac{1}{\beta} (f_n^{KS-2})^{-1} \left(\beta \left[\frac{1}{N} \sum_{k=1}^N m_{(n):k} - m_{\min} \right] \right).$$

The main problem is that the Kijko-Sellevoll functions f_n^{KS-1} and f_n^{KS-2} map the interval $[m_{\min}, m_{\min} + H_n/\beta] \subseteq [m_{\min}, m_{\max}]$ into $[m_{\min}, \infty]$. So if the estimator is greater than the value $m_{\min} + H_n/\beta$ then the solution is beyond infinity. For example let $b=1$, $m_{\max}=8$, $m_{\min}=5$ and $n=1$, then the estimator of the maximum observed magnitude must lie in the interval $[5, 5.4342]$. Due to the maximum magnitude is 8, it is clear that in this case, when $n=1$, often there will be the event greater than 5.4342, and consequently, with mean value superior to this upper limit.

Examples and simulations

In figure 2 we show the problem of simulation for $b=1$, $m_{\max}=8$, $m_{\min}=5$. To each catalogue size (that is to say n) it was generated 1000 artificial catalogues. The figure 2b shows how many of them were accepted i.e. the maximum observed value of catalogue was smaller than $m_{(n):k} < m_{\min} + H_n/\beta$. We calculate the estimator of the maximum for each of the accepted catalogues and afterwards the mean values of those maximums of the catalogues of each size n . The results are plotted in the figure 2a. In the figure 2c we applied the formula (23). For each size n , all the 1000 catalogues are used to estimate the expected value of maximum and by using this estimator (only one value) we calculate the value of the estimator of the maximum \hat{m}_{\max} . The figure 2d shows how many attempts were necessary to make to get a mean value of 1000 catalogues, which fulfills $\bar{m}_{(n)} < m_{\min} + H_n/\beta$. We can see from 2c and 2d that for catalogues of size greater than 10, the method behaves quite stable.

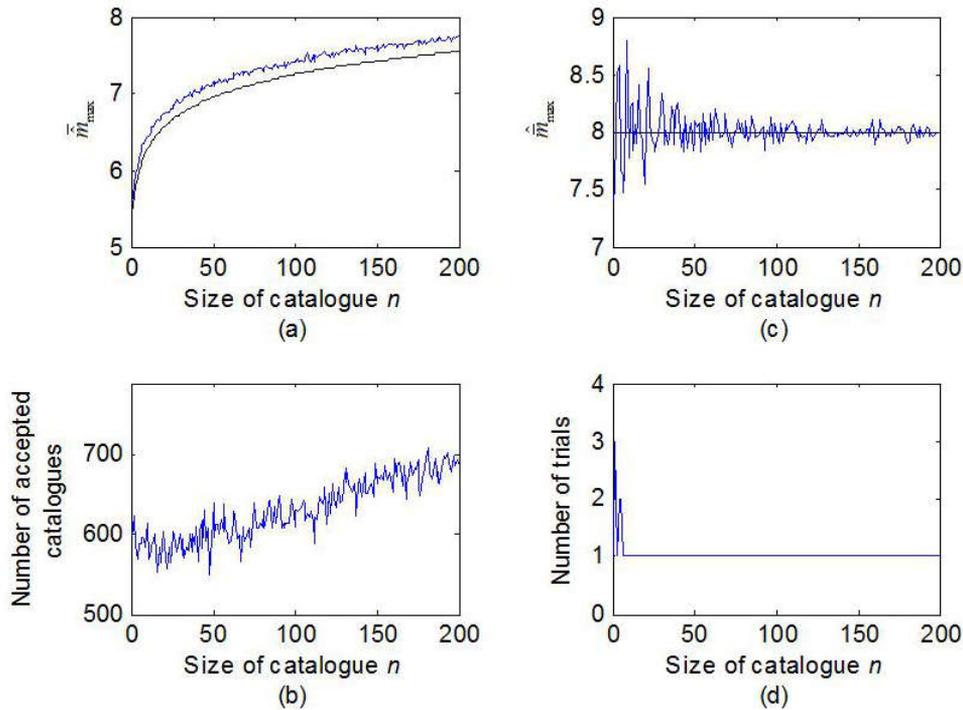


Figure 2: The case $1(8-5)$ and sample size 1000 simulation.

This result is quite expected. In the first case (illustrated by figure 2a) we are modeling the expected value of the maximum with the mean value of only one event so the variance is quite big. For small size catalogues we have 55 to 60% of acceptance of the requirements. The situation is not so much better even though the catalogue size is 200, since still almost 30 % of maximums are rejected. In the second case (figure 2c) all the 1000 maximums are used to calculate the mean value estimator. Its variance decreases as the square root of size of sample, so the variance comes 30 times smaller than in the first case.

Of course we could have enough big catalogue size that all events are included into the acceptable interval, so the maximum m_{\max} (unknown) also does; to the number of events n , holds $m_{\max} \leq m_{\min} + H_n/\beta$. For example if $m_{\min} = 5$ and $b = 1$, then n is 56, 561, 5615 and 56146 for the maximum values 7, 8, 9 and 10, respectively. With the real catalogues this is not possible because we should wait years or hundreds years to gather more data.

Another thing what we can do is to increase the minimum value; for example if $m_{\min} = 6$, then n is, 6, 56, 561 and 5615 for the maximum values 7, 8, 9 and 10, respectively.

Next we consider less number of catalogues. The figure 3 shows the result of considering $m_{\min} = 6$, and $b = 1$ and 100 simulated catalogues; the figure 3b shows that we could use a single catalogue when n is more than about 50. Similarly figure 4 displays the case when minimum value is 7 and only 10 simulated catalogues are taken into account; the figure 4b shows that

almost always it is possible to use one single catalogue if the distance to maximum is about one magnitude unit. Those three figures show that the situation of the single catalogue comes better as the distance between maximum and minimum comes smaller. Also 2c, 3c and 4c shows that it is possible to analyze the behavior of KS function in extreme cases like small catalogue sizes and/or big value of $b(m_{max} - m_{min})$, which can be of extremely importance in zones with few

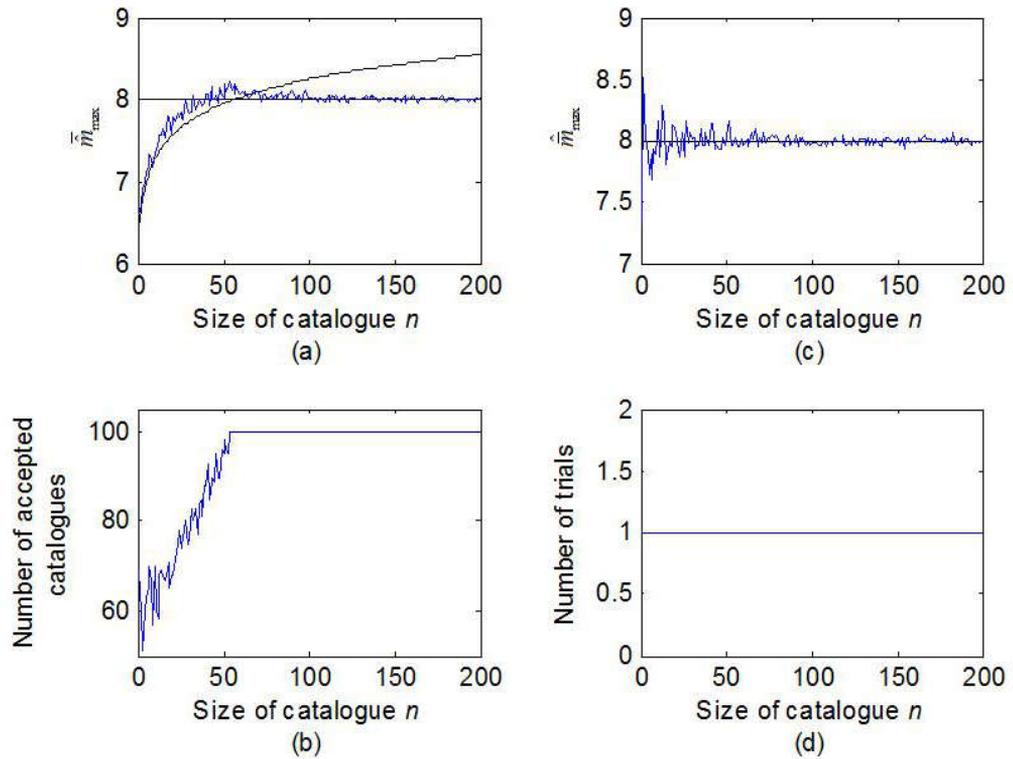


Figure 3: The case 1 (8 - 6) and sample 100 size simulation.

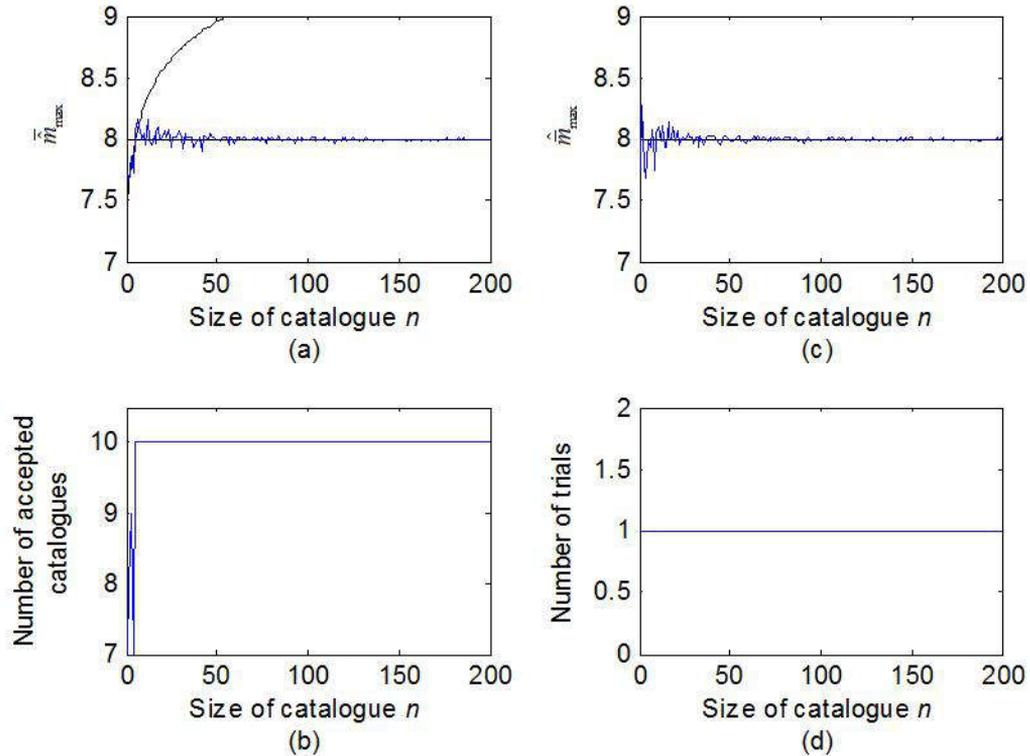


Figure 4: The case 1 (8 - 7) and sample size 10 simulation.

Concluding remarks

In this work we report a method to solve exactly Kijko-Sellevoll formula to calculate $b(m_{\max} - m_{\min})$, considering throughout the text an exact value for b . In fact, both parameters are closely related when we are dealing with seismic hazard assessment. In a following paper we shall show that the estimators of Cosentino et al. (1976, 1977), Page (1968) and Aki-Utsu (1965) are related with the KS-2.

The series resulted to be not only a tool to solve an equation but they also let us to build a rich theory. They give numerically stable method to manage wider range of magnitudes and size of catalogues. The cost of this is to have more complicated calculus (but not so much slower). The exact solution of KS estimator does not only mean the solution of the problem without approximations, besides it makes possible a numerically «exact» solution and the improvement of the computer performance. At least our work gives an alternative viewpoint to see and analyze

other similar methods.

We used a fixed β to all catalogues. In fact this is not realistic since always the β -value must be estimated to each catalogue and that estimator changes from one catalogue to another. That could make the method more «soft,» but still there will be failed catalogues. We shall go insight into this topic in another report (Part II). As we showed, the way to avoid the problem in the case of failing catalogues is to put the minimum m_{min} bigger even the number of the events of the catalogue will come smaller. As we could see from the figures, when the minimum m_{min} is closer to the maximum m_{max} , we need less data to get answers. The variance of the estimator of the maximum comes smaller as the difference $m_{max} - m_{min}$ comes smaller, even we have used fewer events in catalogues. Kijko (2004) showed this fact empirically in his simulations.

Acknowledgments

We are grateful to J. Seppänen for his help to get a valuable paper necessary to complete this work.

Appendix A

To the readers, who are not familiar with KS estimator, we shall give it shortly (the reader can find more details in the works of Kijko and Graham (1998) and Kijko (2004)).

First we remark that Cramér's (1961) approximation

$$[f(x)]^n \approx \exp\{-n[1-f(x)]\}$$

can be derived (using) as

$$\begin{aligned} 0 \leq [f(x)]^n &= \exp\{n \log(f(x))\} \\ &= \exp\{-n[-\log(1-[1-f(x)])]\} \\ &= \exp\left\{-n \sum_{k=1}^{\infty} \frac{[1-f(x)]^k}{k}\right\} \\ &= \exp\left\{-n[1-f(x)] - n \sum_{k=2}^{\infty} \frac{[1-f(x)]^k}{k}\right\} \\ &= \exp\{-n[1-f(x)]\} \exp\left\{-n \sum_{k=2}^{\infty} \frac{[1-f(x)]^k}{k}\right\} \\ &\leq \exp\{-n[1-f(x)]\}. \end{aligned}$$

We see that this approximation comes from the linearization of logarithm and equality holds when $n=0$ or $f(x)=1$. This inequality also shows that Cramér's approximation overestimate the original CDF.

Applying Cramér's approximation to integral (5) we have

$$\begin{aligned} \Delta &\approx \int_{m_{\min}}^{m_{\max}} \exp\left\{-n \left[1 - \frac{1 - \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]}\right]\right\} dm \\ &= \int_{m_{\min}}^{m_{\max}} \exp\left\{-\frac{n \exp(-\beta(m - m_{\min}))}{1 - \exp(-\beta(m_{\max} - m_{\min}))} + \frac{n \exp(-\beta(m_{\max} - m_{\min}))}{1 - \exp(-\beta(m_{\max} - m_{\min}))}\right\} dm \\ &= \frac{\int_{m_{\min}}^{m_{\max}} \exp\left\{-\frac{n \exp(-\beta(m - m_{\min}))}{1 - \exp(-\beta(m_{\max} - m_{\min}))}\right\} dm}{\exp\left\{-\frac{n \exp(-\beta(m_{\max} - m_{\min}))}{1 - \exp(-\beta(m_{\max} - m_{\min}))}\right\}} \\ &=: \end{aligned}$$

Setting now

$$\zeta = \frac{n \exp(-\beta(m - m_{\min}))}{1 - \exp(-\beta(m_{\max} - m_{\min}))} \Rightarrow \frac{d\zeta}{dm} = -\beta\zeta$$

$$n_1(m_{\max}) = \frac{n}{1 - \exp(-\beta(m_{\max} - m_{\min}))}$$

$$n_2(m_{\max}) = n_1(m_{\max}) \exp(-\beta(m_{\max} - m_{\min}))$$

we can write

$$\therefore = \frac{- \int_{\frac{n_1(m_{\max})}{n_2(m_{\max})}}^{\frac{n_2(m_{\max})}{n_1(m_{\max})}} \exp(-\zeta) \frac{d\zeta}{\beta\zeta}}{\exp(-n_2(m_{\max}))} = \frac{\int_{n_2(m_{\max})}^{\infty} \frac{\exp(-\zeta)}{\zeta} d\zeta - \int_{n_1(m_{\max})}^{\infty} \frac{\exp(-\zeta)}{\zeta} d\zeta}{\beta \exp(-n_2(m_{\max}))}.$$

Note that $n_1(\mathfrak{M}) - n_2(\mathfrak{M}) = n$ for all \mathfrak{M} . We can approximate the exponential integral by

$$E_1(z) = \int_z^{\infty} \frac{\exp(-\zeta)}{\zeta} d\zeta = \frac{z^2 + a_1z + a_2}{z(z^2 + b_1z + b_2)} \exp(-z).$$

$$a_1 = 2.334733 \quad a_2 = 0.250621$$

$$b_1 = 3.330657 \quad b_2 = 1.681534$$

KS estimator is given now

$$\mathfrak{M} = m_{(n)} + \frac{E_1(n_2(\mathfrak{M})) - E_1(n_1(\mathfrak{M}))}{\beta \exp(-n_2(\mathfrak{M}))}. \quad (24)$$

This also can be expressed as

$$\hat{m}_{\max} = m_{(n)} + \frac{E_1(n_2(m_{(n)})) - E_1(n_1(m_{(n)}))}{\beta \exp(-n_2(m_{(n)}))}. \quad (25)$$

Kijko (2004) remarked that this estimator (25) can be used when $(m_{\max} - m_{\min}) < 2$ and $n > 100$.

Appendix B

In this section we discuss the numerical solution of series. The exact solution of the integral is the «core» of the Newton-Raphson method. For example with $\beta(m_{\max} - m_{\min}) = 1$ and $n = 400$ the solution (9) comes unstable. This is clear because we showed that the numerator is a tail of logarithm function (so the numerator is small as n is big) and, if at the same time $1 - \exp(-\beta(m_{\max} - m_{\min})) \ll 1$ the denominator is very small, yielding equation (9) to instability.

Because of all the series of the Kijko-Sellevoll functions are nonnegative terms series, they have not a similar numerical instability. Because of $\{a_k\}$ is nonnegative sequence, we could calculate the series $S = \sum_{k=0}^{\infty} a_k$ as a partial sum $S_{n_\varepsilon} = \sum_{k=0}^{n_\varepsilon} a_k$, where n_ε is some integer such as $S_{n_\varepsilon} = S_n$ when $n \geq n_\varepsilon$.

The idea sounds quite simple, but the «coin has also other side». The Kijko-Sellevoll function KS-1 belongs to the family of Lerch transcendent function (or shortly Lerch Phi) (Lerch, 1887; Erdélyi et al, 1953))

$$\Phi(z, s, \alpha) = \sum_{k=0}^{\infty} \frac{z^k}{(k + \alpha)^s}$$

and they both (KS-1 and KS-2) are close to logarithm function (we showed that KS-1 is scaled tail of logarithm function). This means that the convergence of the series is slowly, so we need to use some acceleration algorithm to solve the value of the series. For example in the extreme case when $b(m_{\max} - m_{\min}) = 16$ we might need approximately 10^{19} terms to calculate the value using the direct sum of the series, but with accelerator (discussed below) we need only about 10^3 terms from the series.

The following algorithm we present here, bases on the paper of Cohen et al. (2000) and their algorithm \mathcal{Z}_A . We tried also the algorithm \mathcal{Z}_B , but we did not get better results than those reached by the former expression, so we adopted it, even Cohen et al. (2000) recommended method \mathcal{Z}_B . This algorithm presented below can be applied also to other alternative or nonnegative series different than KS functions. Because of the numerical solution of the integral is so important tool to the analysis of the exact solution of Kijko-Sellevoll estimator, we give also the open source code in MATLAB.

Having a series

$$S = \sum_{k=0}^{\infty} (-1)^k a_k,$$

where a_k is well-behaved function which goes to zero as $k \rightarrow \infty$ (sequence of the series), we want to find coefficients $c_{n,k}/d_n$ such that the sequence

$$S_n = \sum_{k=0}^{n-1} \frac{c_{n,k}}{d_n} a_k \tag{26}$$

converges quickly to zero i.e. $|S - S_n| < C^{-n}$ to some constant C . Cohen et al (2000) showed that algorithm \mathcal{Z}_A has convergence factor 7.89 for a large class of sequences $\{a_k\}$ and 17.93 for a small class of sequences. To the algorithm \mathcal{Z}_B was reported factor 9.56 for a large class and 14.41

for a small class. From our experiments, we can say that our sequences converge approximately with factor 10 which means a one correct decimal in the sequence (26) for each term in the sum.

The algorithm of Cohen et al. bases on Chebyshev polynomials. It is set $P_n(\sin^2 t) = \cos 2nt$ so that $P_n(x) = T_n(1-2x)$, where $T_n(x)$ is the ordinary Chebyshev polynomial. Clearly $P_0(x) = 1$. Any way this can be any arbitrary constant. Now the sequence of polynomials can be given as

$$P_n(x) = \sum_{m=0}^n (-1)^m \frac{n}{n+m} \binom{n+m}{2m} 2^{2m} x^m.$$

Using these polynomials, it can be defined a new family of polynomials

$$P_n^{(m)}(x) = \sum_{r=0}^m (-1)^r \binom{m}{r} (n-2r)^{m+1} P_{n-2r}(x).$$

Here we can see that when $n-2r = 0$ (case when we have P_0) the factor of the polynomial is zero, so P_0 can be any constant. The suggested sequence of polynomials are defined as

$$A_n(x) = \frac{P_n^{(n-1)}(x)}{n! 2^{n-1}}.$$

The normalization factor has been chosen to fulfill $A_n(0) = 1$. Any way it is an arbitrary factor and we could choose directly $A_n(x) = P_n^{(n-1)}(x)$. The algorithm of Cohen et al (2000) is

$$\begin{aligned} \text{Let } A_n(x) &= \sum_{k=0}^n b_k x^k \\ d &= A_n(-1); c_{n,0} = -d; s = 0 \\ \text{For } k &= 0 \text{ up to } k = n-1, \text{ repeat:} \\ c_{n,k+1} &= -b_k - c_{n,k}; \quad s = s + c_{n,k+1} \cdot a_k; \\ \text{Output: } & s/d \end{aligned}$$

This algorithm has been written to evaluate the factors «on the fly.» In our case the degree of the polynomial is fixed because we do not want to use time to recalculate the factors in each time when the program is called. These factors are universals and their values depend only on the degree of the polynomial. From the algorithm we can see that the factors are

$$\begin{aligned} c_{n,0} &= -\sum_{m=0}^n |b_m|, \quad c_{n,1} = \sum_{m=1}^n |b_m|, \quad c_{n,2} = -\sum_{m=2}^n |b_m|, \quad \dots \\ c_{n,k} &= (-1)^{k+1} \sum_{m=k}^n |b_m|, \end{aligned}$$

so $d = -c_{n,0}$. Normalized factors result equal to $\tilde{c}_{n,k} = c_{n,k}/d = -c_{n,k}/c_{n,0}$, which in the case of $n = 18$, give the next values:

$$\begin{aligned}
 \tilde{c}_{18,1} &= 0.9999999999999245 & \tilde{c}_{18,7} &= 0.999816879032895474 & \tilde{c}_{18,13} &= 0.580476889354509827 \\
 \tilde{c}_{18,2} &= -0.99999999998277922 & \tilde{c}_{18,8} &= -0.998595555793887371 & \tilde{c}_{18,14} &= -0.356402292890562931 \\
 \tilde{c}_{18,3} &= 0.99999999610062291 & \tilde{c}_{18,9} &= 0.992342806734044361 & \tilde{c}_{18,15} &= 0.168952372776841569 \\
 \tilde{c}_{18,4} &= -0.99999972513434920 & \tilde{c}_{18,10} &= -0.969204555863406323 & \tilde{c}_{18,16} &= -0.057152687879365596 \\
 \tilde{c}_{18,5} &= 0.999999109387382570 & \tilde{c}_{18,11} &= 0.906022499505876388 & \tilde{c}_{18,17} &= 0.012162740075262281 \\
 \tilde{c}_{18,6} &= -0.999983872560606631 & \tilde{c}_{18,12} &= -0.777218084969806462 & \tilde{c}_{18,18} &= -0.001216274007526228
 \end{aligned}$$

Pay attention that in the MATLAB code the factors are not normalized.

We can also see from the algorithm that the partial sum is now

$$S = -\frac{1}{c_0} \sum_{k=0}^{n-1} c_{n,k+1} a_k = \sum_{k=0}^{n-1} \tilde{c}_{n,k+1} a_k.$$

The algorithm above is to the alternative series. The nonnegative series can be solved by means of the trick of Van Wijngaarden (Press et al., 1992)

$$\sum_{k=1}^{\infty} a_k = \sum_{m=1}^{\infty} (-1)^m b_m \quad \text{with} \quad b_m = \sum_{k=0}^{\infty} 2^k a_{2^k m}.$$

```

function S = KS(ftype,x,n)
%
% Input:
% ftype = 1: Kijko-Sellevoll function 1
%         2: Kijko-Sellevoll function 2
%         3: A special series for a variance
% x      = beta*(mmax-mmin) (scalar or vector)
% n      = number of events (scalar or vector)
%
% Written by Mika Haarala Orosco, Acrenet Oy
% (21.1.2015 - ver. 14.06.2016)
%
% Reference:
% Cohen, H., F. Rodriguez Villegas, and D. Zagier (2000). Convergence

```

```
% Cohen, H., F. Rodriguez Villegas, and D. Zagier (2000). Convergence
% acceleration of alternating series, Exper. Math. 9, 3-12.

miter = 10000;
if ~(ftype == 1 || ftype == 2 || ftype == 3), error('Ftype must be 1, 2 or 3. '),
end
if ~(all(size(x) == size(n)) || numel(x) == 1 || numel(n) == 1)
    error('Inputs must be vectors or scalars.')
end

S = NaN( max( size(x), size(n) ) );
z = 1 - exp(-x(:));
n = n(:);

if ftype == 1
    I = z == 1;
    S(I) = Inf;
else
    I = false( size(z) );
end

I = ~(z < 0 | I | n < 0);
if ~isscalar(z)
    z = z(I);
end
if ~isscalar(n)
    n = n(I);
end

Sn = zeros( size(z) );
So = -1;

if any(z > 0.35)
```

```

% Accelerated sum

f=[1.437775728963973375
-1.437775728961498499
 1.437775728403331487
-1.437775689444458316
 1.437774448462769210
-1.437752541323044337
 1.437512442082007163
-1.435756453171741645
 1.426766402334197056
-1.393498786821714133
 1.302657159684823614
-1.117465298681447728
 0.834595582738420712
-0.512426566465161050
 0.242915620929416540
-0.082172747478005376
 0.017487292477909572
-0.001748729247790957];

for k = 1:length(f)
    fi = 1;
    for i = 1:miter
        switch ftype
            case 1
                Sn = Sn + f(k)*(fi .* z.^(fi*k)./(fi*k + n));
            case 2
                Sn = Sn + f(k)*(fi .* n .* z.^(fi*k)./(fi*k * (fi*k + n)));
            case 3
                Sk = zeros(size(n));
                fik= fi*k;

```

```

        for j = 1:length(n)
            if fik < 3000
                Sk(j) = sum( 1./ ( n(j)+1 : n(j)+fik-1 ) );
            else
                Sk(j) = Hn( n(j), n(j)+fik-1 );
            end
        end
        Sn = Sn + f(k)*(fi * 2*n .* Sk .* z.^(fik + 1) ./ ...
            ((2*n + fik) .* (fik + 1 + n)) );
    end
    if all(Sn == So), break, end
    So = Sn;
    fi = fi*2;
end
end
Sn = Sn/1.437775728963974460;

else
    % Direct sum

    for k=1:miter
        switch ftype
            case 1
                Sn = Sn + z.^k ./ (k + n);
            case 2
                Sn = Sn + n .* z.^k ./ (k *(k + n));
            case 3
                Sk = zeros(size(n));
                for j = 1:length(n)
                    Sk(j) = sum( 1./ ( n(j)+1 : n(j)+k-1 ) );
                end
                Sn = Sn + 2*n .* Sk .* z.^(k+1) ./ ( (2*n + k) .* (k + n) );
            end
        end
    end
end

```

```

        if all(Sn == So), break, end
        So = Sn;
    end

end

S(I) = Sn;

function y = Hn(k1,k2)
%
% This function calculates the subtraction of Harmonic numbers:
%   y = sum(1./(k1+1:k2)) = sum(1./(1:k2)) - sum(1./(1:k1))
%
% Reference:
%   Villarino, M.B. 2008, Ramanujan's Harmonic Number Expansion
%   into Negative Powers, J. Inequal. Pure and Appl. Math., 9(3),
%   Art. 89, 12 pp.
%
euler = 0.57721566490153286;

if k1 < 10
    D = ceil(10-k1);
    m = (k1 + D) * (k1 + D + 1) / 2;
    y1 = euler + log(2*m)/2 + 1/(12*m) - 1/(120*m^2)...
        + 1/(630*m^3) - 1/(1680*m^4) + 1/(2310*m^5)...
        - 191/(360360*m^6) - sum( 1./ (k1+D:-1:k1+1) );
else
    m = k1 * (k1+1) / 2;
    y1 = euler + log(2*m)/2 + 1/(12*m) - 1/(120*m^2)...
        + 1/(630*m^3) - 1/(1680*m^4) + 1/(2310*m^5)...

```

end

$m = k2 * (k2+1) / 2;$

$y2 = \text{euler} + \log(2*m) / 2 + 1 / (12*m) - 1 / (120*m^2) \dots$
 $+ 1 / (630*m^3) - 1 / (1680*m^4) + 1 / (2310*m^5) \dots$
 $- 191 / (360360*m^6);$

$y = y2 - y1;$

References

- Abramowitz, M., and I. A. Stegun, «Handbook of mathematical functions», 10th ed., Dover Publ., New York, 1972.
- Aki, K., «Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits», *Bull. Earthquake Res. Inst. Tokyo Univ.*, vol. 43, pages 237-239, 1965.
- Amorèse, D., J.-R. Grasso, and P. A. Rydelek, «On varying b -values with depth: results from computer-intensive tests for Southern California», *Geophys. J. Int.*, vol. 180, pages 347-360, 2010.
- Anagnostopoulos, S., C. Providakis, P. Salvaneschi, G. Athanasopoulos, and G. Bocacina, «SEISMOCARE: An efficient GIS tool for scenario-type investigations of seismic risk of existing cities», *Soil Dyn. Earthquake Eng.*, vol. 28, pages 73-84, 2008.
- Cohen, H., F. Rodriguez Villegas, and D. Zagier, «Convergence acceleration of alternating series», *Exper. Math.*, vol. 9, pages 3-12, 2000.
- Cosentino, P., and D. Luzio, «A generalization of the frequency-magnitude relation in the hypothesis of a maximum regional magnitude», *Ann. Geofis. (Rome)*, vol. 29, 1-2, pages 3-8, 1976.
- Cosentino, P., V. Ficara, and D. Luzio, «Truncated exponential frequency-magnitude relationship in the earthquake statistics», *Bull. Seism. Soc. Am.*, vol. 67, pages 1615-1623, 1977.
- Cramér, H., «Mathematical methods of statistics», 9th ed., Princeton University Press, Princeton, 1961.
- Erdélyi, A., W. Magnus, F. Oberhettinger, and F. G. Tricomi, «Higher transcendental functions», vol. 1, McGraw-Hill Book Company, Inc., New York, 1953.

- Gutenberg, B., and C. F. Richter, «Frequency of earthquakes in California», *Bull. Seism. Soc. Am.*, vol. 34, pages 185-188, 1944.
- Hamilton, R. M., «Mean magnitude of an earthquake sequence», *Bull. Seism. Soc. Am.*, vol. 57, pages 1115-1116, 1967.
- Holschneider, M., G. Zöller, and S. Hainzl, «Estimation of the maximum possible magnitude in the framework of a doubly truncated Gutenberg-Richter model», *Bull. Seism. Soc. Am.*, vol. 101, pages 1649-1659, 2011.
- Ishibe, T., and K. Shimazaki, «The Gutenberg-Richter relationship vs. the characteristic earthquake model: Effects of different sampling methods», *Bull. Earthq. Res. Inst Tokyo Univ.*, vol. 83, pages 131-151, 2008.
- Ishimoto, M., and K. Iida, «Observations of earthquakes registered with the microseismograph constructed recently», *Bull. Earthquake Res. Inst.*, vol. 17, pages 443-478, 1939.
- Lerch, M., «Note sur la fonction $\mathfrak{K}(w, x, s) = \sum_{k=0}^{\infty} \exp(2k\pi ix) / (w+k)^s$ », *Acta Math.*, vol. 11, 1-4, pages 19-24, 1887.
- Kahraman, S., T. Baran, Ý. A. Saatçi, and M. Palk, «The effect of regional borders when using the Gutenberg-Richter model, case study: Western Anatolia», *Pure Appl. Geophys.*, vol. 165, pages 331-347, 2008.
- Kijko, A., «Is it necessary to construct empirical distributions of maximum earthquake magnitudes?», *Bull. Seism. Soc. Am.*, vol. 74, pages 339-347, 1984.
- Kijko, A., and M. A. Sellevoll, «Estimation of earthquake hazard parameters from incomplete data files. Part I: Utilization of extreme and complete catalogues with different threshold magnitudes», *Bull. Seism. Soc. Am.*, vol. 79, pages 645-654, 1989.
- Kijko, A., and G. Graham, «'Parametric-historic' procedure for probabilistic seismic hazard analysis. Part I: Assessment of maximum regional magnitude m_{\max} », *Pure Appl. Geophys.*, vol. 152, pages 413-442, 1998.
- Kijko, A., «Estimation of the maximum earthquake magnitude, m_{\max} », *Pure Appl. Geophys.*, vol. 161, pages 1-27, 2004.
- Kijko, A., and A. Smit, «Extension of the Aki-Utsu b -value estimator for incomplete catalogs», *Bull. Seism. Soc. Am.*, vol. 102, pages 1283-1287, 2012.
- Leyton, F., S. Ruiz, and S. A. Sepúlveda, «Preliminary re-evaluation of probabilistic seismic hazard assessment in Chile: from Arica to Taitao Peninsula», *Av. Geosci.*, vol. 22, pages 147-153, 2009.

- Márquez-Ramírez, V. H., F. A. Nava, and F. R. Zúñiga, «Correcting the Gutenberg-Richter b -value for effects of rounding and noise», *Earthq. Sci.*, vol. 28, pages 129-134, 2015.
- Mostafanejad, A., C. A. Powell, and C. A. Langston, «Variation of seismic b -value in the New Madrid seismic zone: Evidence that the Northern Reelfoot fault is creeping», *Seismol. Res. Lett.*, vol. 84, pages 1124-1129, 2013.
- Page, R., «Aftershocks and microaftershocks of the great Alaska earthquake of 1964», *Bull. Seism. Soc. Am.*, vol. 58, pages 1131-1168, 1968.
- Pisarenko, V. F., A. A. Lyubushin, V. B. Lysenko, and T. V. Golubieva, «Statistical estimation of seismic hazard parameters: Maximum possible magnitude and related parameters», *Bull. Seism. Soc. Am.*, vol. 86, pages 691-700, 1996.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, «Numerical recipes in FORTRAN: The art of scientific computing», 2nd ed., Cambridge University Press, Cambridge, 1992.
- Rong, Y. D., D. Jackson, H. Magistrale, and C. Goldfinger, «Magnitude limits of subduction zone earthquakes», *Bull. Seism. Soc. Am.*, vol. 104, pages 2359-2377, 2014.
- Utsu, T., «A method for determining the value of b in a formula $\log n = a - bM$ showing the magnitude-frequency relation for earthquakes», *Geophys. Bull. Hokkaido Univ.*, vol. 13, pages 99-113, 1965.
- Villarino, M.B., «Ramanujan's Harmonic Number Expansion into Negative Powers», *J. Inequal. Pure and Appl. Math.*, vol. 9 (3), Art. 89, 12 pp, 2008.
- Zúñiga, F. R., and A. Figueroa-Soto, «Converting magnitudes based on the temporal stability of a - and b -values in the Gutenberg-Richter law», *Bull. Seism. Soc. Am.*, vol. 102, pages 2116-2127, 2012.

Analysis of Gutenberg-Richter b -value and m_{\max} Part II: Estimators for b -value

M. Haarala^{1,2}, L. Orosco^{1,3}

Abstract

This report is the second of a series of three which have the main goal to achieve a method to estimate the parameters b and m_{\max} that are essential when Gutenberg-Richter law is used for seismic hazard assessment. This paper is devoted to analyze the estimators of b -value.

We give an estimator for the expected value of $M_{(N)}$ and an exact and numerically stable variance for it; $M_{(N)}$ represents the ordered values (by size) of magnitude catalogue.

We go on the development of the theory of Kijko-Sellevoll functions, applying them to calculate generalized b estimators of the seismic catalogue.

Keywords: m_{\max} , b -value, Gutenberg-Richter distribution function - Kijko Sellevoll estimator

Resumen

Este artículo es el segundo de una serie de tres, que tienen como objetivo principal obtener un método para calcular los parámetros b y m_{\max} que son fundamentales cuando se utiliza la ley de Gutenberg – Richter para la estimación de la peligrosidad sísmica. En este trabajo en particular, nos enfocamos al análisis de estimadores del valor b .

Proponemos un estimador para el valor esperado de $M_{(N)}$ y una expresión exacta y numéricamente estable para su varianza; $M_{(N)}$ representa los valores ordenados de las magnitudes del catálogo sísmico.

Continuamos con el desarrollo de la teoría de las funciones de Kijko-Sellevoll, aplicándolas al cálculo de estimadores de b generales.

Palabras clave: m_{\max} - b - función de distribución de Gutenberg-Richter - estimador de Kijko Sellevoll

Introduction

We are analyzing the double truncated Gutenberg-Richter distribution function

¹ Instituto de Estudios Interdisciplinarios de Ingeniería (IESIING) – Facultad de Ingeniería – UCASAL

² Acrenet Oy, Joensuu, Finland

³ Facultad de Ingeniería – U.N.Sa.

$$f(m) = \frac{\beta \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]},$$

which has cumulative distribution function (CDF)

$$F_M(m | m_{\max}) = \begin{cases} 0, & \text{for } m < m_{\min}, \\ \frac{1 - \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]}, & \text{for } m_{\min} \leq m < m_{\max}, \\ 1, & \text{for } m_{\max} \leq m. \end{cases} \quad (1)$$

The limit distribution function was given as

$$F_M(m | m_{\max} = \infty) = \begin{cases} 0, & \text{for } m < m_{\min}, \\ 1 - \exp[-\beta(m - m_{\min})], & \text{for } m_{\min} \leq m, \end{cases}$$

and

$$F_M(m | m_{\max} = m_{\min}) = \begin{cases} 0, & \text{for } m < m_{\min}, \\ 1, & \text{for } m_{\min} \leq m, \end{cases}$$

where $\beta = b \log(10)$, m_{\min} is a threshold magnitude and m_{\max} is a maximum possible magnitude. We could assume that $m_{\max} \in [m_{\min}, \infty]$. If we let $M_1, M_2, \dots, M_N \in [m_{\min}, m_{\max}]$ be a set of random variables from the catalogue C_N of size N , and we let $M_{(1)} \leq M_{(2)} \leq \dots \leq M_{(N)}$ denote the ordered values of M_1, M_2, \dots, M_N , so the random variable $M_{(N)}$ is a maximum in the catalogue C_N . We assume that these random variables are independently and identically distributed (iid) with CDF of $F_M(m)$ given by (1). Let now $m_{(1)} \leq m_{(2)} \leq \dots \leq m_{(N)}$ to be an ordered sample of magnitudes, where $m_{(1)}$ is a minimum observed magnitude ($m_{\min} \leq m_{(1)}$) and $m_{(N)}$ is a maximum observed magnitude ($m_{(N)} \leq m_{\max}$), having a CDF

$$F_{M_{(n)}}(m | m_{\max}) = \begin{cases} 0, & \text{for } m < m_{\min}, \\ [F_M(m | m_{\max})]^n, & \text{for } m_{\min} \leq m < m_{\max}, \\ 1, & \text{for } m_{\max} \leq m. \end{cases} \quad (2)$$

We showed in a previous paper (Haarala and Orosco, 2016) that using a Kijko-Sellevoll function 1 (KS-1)

$$f_n^{KS-1}(x) = \sum_{k=1}^{\infty} \frac{(1 - \exp[-x])^k}{k + n}$$

or a Kijko-Sellevoll function 2 (KS-2)

$$f_n^{KS-2}(x) = n \sum_{k=1}^{\infty} \frac{(1 - \exp[-x])^k}{k(k + n)}$$

we can write the expected value of the maximum $M_{(n)}$ of the catalogue as

$$\begin{aligned} E(M_{(n)} | m_{max}) &= m_{max} - \frac{1}{\beta} f_n^{KS-1}(\beta(m_{max} - m_{min})) \\ &= m_{min} + \frac{1}{\beta} f_n^{KS-2}(\beta(m_{max} - m_{min})). \end{aligned} \quad (3)$$

The relation between KS-1 and KS-2 functions is

$$f_n^{KS-1}(\beta(m_{max} - m_{min})) + f_n^{KS-2}(\beta(m_{max} - m_{min})) = \beta(m_{max} - m_{min}). \quad (4)$$

Estimator for expected value $M_{(N)}$

Suppose that we have a set of iid events m_1, \dots, m_N from the catalogue C_N . We can divide C_N into N sub-catalogues such that each sub-catalogue $C_{1:k}$ (where $k = 1, \dots, N$) has one and only one event. Each of them makes a catalogue of size $n = 1$ with maximum observed magnitude $m_{(1):k}$. Because of each catalogue has only one event, $m_{(1):k} = m_k$. The mean value of maximum observed values is the unbiased estimator for the expected value $M_{(1)}$, thus

$$\hat{E}(M_{(1)} | m_{max}) = \frac{m_{(1):1} + \dots + m_{(1):N}}{N} = \bar{m}_{(1)} = \frac{m_1 + \dots + m_N}{N} = \bar{m}. \quad (5)$$

This shows that the mean value of the maximums of the catalogues $C_{1:k}$ is an estimator for the expected value of $M_{(1)}$ ($M_{(1)}$ is a minimum but also a maximum since each catalogue $C_{1:k}$ has only one event). In the same way we could create the estimator for sub-catalogues $C_{2:k}$ (size $n = 2$; each sub-catalogue has two and only two events)

$$\hat{E}(M_{(2)} | m_{max}) = \frac{m_{(2):1} + \dots + m_{(2):\lfloor N/2 \rfloor}}{\lfloor N/2 \rfloor} = \bar{m}_{(2)},$$

where $\lfloor \cdot \rfloor$ is a floor function (maximum integer value m such that $m \leq N/2$). If the size of catalogue N is odd, then there is a value which does not belong to any sub-catalogue and we can choose it randomly.

In general

$$\hat{E}\left(M_{(n)} \mid m_{\max}\right) = \frac{m_{(n):1} + \dots + m_{(n):\lfloor N/n \rfloor}}{\lfloor N/n \rfloor} = \bar{m}_{(n)}, \quad (6)$$

where $1 \leq n \leq N$ and $m_{(n):k}$ is a maximum of the sub-catalogue $C_{n:k}$ ($1 \leq k \leq \lfloor N/n \rfloor$), which has n and only n events. The estimator in the case is $n = N$ is

$$\hat{E}\left(M_{(N)} \mid m_{\max}\right) = \frac{m_{(N):1}}{1} = \bar{m}_{(N)} = m_{(N)}. \quad (7)$$

The mean value from one event is the event itself. This shows also that the maximum observed value is the unbiased estimator for the expected value of maximum of the catalogue C_N . Pisarenko et al (1996) showed that (7) is the best unbiased estimator and equation (6) yields to (7) when the sub-catalogue size is the same than catalogue size itself.

In the case that $n = 1$, each sub-catalogue has only one event which is also the maximum so we will always get the same mean, independently the way we select the events from the catalogue.

When $n = N$, we have only one sub-catalogue (which has the same events than the catalogue), so we will have always the same maximum, indistinctly the order we take the events from catalogue to conform the sub-catalogue,

In the cases $1 < n < N$ the situation is different. For example, if we want to have sub-catalogue $N - 1$, we choose randomly the events from the catalogue C_N . It could happen that the sub-catalogue $C_{N-1:1}$ has or has not the event $m_{(N)}$. If the event belongs to the sub-catalogue, its estimator for the maximum is $m_{(N)}$, other way it is $m_{(N-1)}$. This situation will present for all n which is not 1 or N . This is the reason for choosing the events randomly to create the sub-catalogues. Of course the simulated catalogue is random and it is not necessarily to randomize the events, but randomizing the selected events we can get different values for the estimator. Real catalogues can have some systematic changes of the β -value (see for example Cao and Gao, 2002), and to avoid the bias, what can happen by rejecting systematically the last terms of the catalogue, it is better to choose the sub-catalogues randomly.

In a previous paper (Haarala and Orosco, 2016) we used the same technique to find an estimator for the expected value. There we generated 1000 artificial catalogues of size n . From each one we took the maximum and calculated the mean of them. In that case we had unbounded number of total events. In the case of estimator (6) the total number of events is bounded (like it is in earthquake catalogue) and the number of the maximums in the mean value depends upon the number of the sub-catalogues, what is possible to get from those limited number of events.

Estimators for β -value

In the case we consider the sub-catalogues of size $n = 1$ we can write the expected values (3) as

$$\bar{m}_{(1)} = \bar{m} = m_{\max} - \frac{1}{\beta} f_1^{KS-1}(\beta(m_{\max} - m_{\min}))$$

in case KS-1 and

$$\bar{m}_{(1)} = \bar{m} = m_{\min} + \frac{1}{\beta} f_1^{KS-2}(\beta(m_{\max} - m_{\min})).$$

in case KS-2. The KS-1 function can write in the form of a partial sum of $n = 1$ (Haarala and Oroscio, 2016)

$$\begin{aligned} \bar{m} &= m_{\max} - \frac{1}{\beta} \frac{\beta(m_{\max} - m_{\min}) - (1 - \exp[-\beta(m_{\max} - m_{\min})])}{(1 - \exp[-\beta(m_{\max} - m_{\min})])} \\ &= \frac{m_{\min} - m_{\max} \exp[-\beta(m_{\max} - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} + \frac{1}{\beta}. \end{aligned} \quad (8)$$

The same result was given also by Hamilton (1968), in the form

$$\bar{m} = \frac{1}{\beta} + \frac{m_{\max} \exp[-\beta m_{\max}] - m_{\min} \exp[-\beta m_{\min}]}{\exp[-\beta m_{\max}] - \exp[-\beta m_{\min}]} \quad (9)$$

and Cosentino et al. (1976; 1977) as

$$\begin{aligned} \bar{m} &= \frac{1}{\beta} + m_{\max} + \frac{m_{\max} - m_{\min}}{\exp[-\beta(m_{\max} - m_{\min})] - 1} \quad (\text{ver. 1976}) \\ &= \frac{1}{\beta} + m_{\min} + \frac{m_{\max} - m_{\min}}{1 - \exp[\beta(m_{\max} - m_{\min})]}. \quad (\text{ver. 1977}) \end{aligned} \quad (10)$$

All these three (or four) solutions (8)-(10) are equal and evaluated using moment method. Cosentino et al (1976, 1977) used this solution with variance to calculate the estimators for the maximum m_{\max} and the β -value.

The equation (8) gives also the same estimator for β that was given by Page (1968)

$$\hat{\beta}_P = \left[\bar{m} - \frac{m_{\min} - m_{(N)} \exp(-\hat{\beta}_P(m_{(N)} - m_{\min}))}{1 - \exp(-\hat{\beta}_P(m_{(N)} - m_{\min}))} \right]^{-1} \quad (11)$$

where $m_{(N)}$ is a maximum observed value in catalogue C_N . Page derived his estimator using maximum likelihood method. We shall call this (11) estimator as a Page' estimator even the same is found by Hamilton and Cosentino et al. Also we can see from (8) that when $m_{\max} \rightarrow \infty$ then we have Aki-Utsu estimator $\beta_{AU}(\bar{m} - m_{\min}) = 1$. So this estimator (11) is related with the KS-2 function.

It was natural to wait that the maximum or extreme value model of the order statistic gives the same estimator than other models since the distribution functions are the same in case $n = 1$.

Owing to the relation between KS-1 and KS-2 given by (4), we can always easily change from the function to other. So there exists other variation of Page' estimator

$$\begin{aligned} \bar{m} &= m_{\max} - \frac{1}{\beta} \left\{ \beta(m_{\max} - m_{\min}) - \frac{\beta(m_{\max} - m_{\min}) - (1 - \exp[-\beta(m_{\max} - m_{\min})])}{(1 - \exp[-\beta(m_{\max} - m_{\min})])} \right\} \\ &= \frac{m_{\max} - m_{\min} \exp[-\beta(m_{\max} - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} - \frac{1}{\beta}. \end{aligned} \tag{12}$$

Thus

$$\hat{\beta}_P = \left[\frac{m_{(N)} - m_{\min} \exp(-\hat{\beta}_P(m_{(N)} - m_{\min}))}{1 - \exp(-\hat{\beta}_P(m_{(N)} - m_{\min}))} - \bar{m} \right]^{-1} \tag{13}$$

is other form of the Page' estimator and gives exactly same estimates than Page' original estimator. The estimator (12) does not exist as $m_{\max} \rightarrow \infty$ (it is valid only when $m_{\max} < \infty$) so this estimator(13) is related with the KS-1 function.

We showed before (Haarala and Orosco, 2016) that in unbounded case ($m_{\max} = \infty$) the KS-1 is unbounded and the KS-2 gives

$$E(M_{(n)} | \infty) - m_{\min} = \frac{H_n}{\beta}, \tag{14}$$

where the $H_n = \sum_{k=1}^n k^{-1}$ is a harmonic number. Setting $n = 1$ and using the mean value estimator we have the Aki-Utsu estimator $\hat{\beta}_{AU}$ (Aki, 1965, and Utsu, 1965)

$$\hat{\beta}_{AU} = \frac{1}{\bar{m} - m_{\min}}.$$

Because in (14) H_n/β is a constant, it means that the distance between expected and minimum values does not depend how we choose the minimum. Always the distance from the arbitrary minimum to the expected value $E(M_{(n)} | \infty)$ is the same, when $m_{\max} = \infty$. This explain the bias what we can see in the Aki-Utsu estimator as $m_{\min} \rightarrow m_{\max}$. The mean \bar{m} does not follow the minimum similar way than the expected value $E(M_{(n)} | \infty)$.

When the data comes from bounded system, then it is $\hat{E}(M_1 | m_{\max} < \infty) = \bar{m}$. Using KS-2 function

to estimator we get

$$\frac{1}{\hat{\beta}_{AU}} = \bar{m} - m_{\min} = \frac{1}{\hat{\beta}} f_1^{KS-2} \left(\hat{\beta} (m_{\max} - m_{\min}) \right).$$

The relation

$$\hat{\beta} = \hat{\beta}_{AU} f_1^{KS-2} \left(\hat{\beta} (m_{\max} - m_{\min}) \right) \leq \hat{\beta}_{AU} \quad (15)$$

shows that Aki-Utsu estimator overestimate since $\hat{\beta} \leq \hat{\beta}_{AU}$ (equality holds only when $m_{\max} = \infty$). This gives also the correction function for the Aki-Utsu estimator

$$\hat{\beta} \approx \hat{\beta}_{AU} f_1^{KS-2} \left(\hat{\beta}_{AU} (\hat{m}_{\max} - m_{\min}) \right),$$

where the estimator \hat{m}_{\max} is some estimator for m_{\max} .

Similarly, from (15) we can see that the Page's estimator and Aki-Utsu estimator has a relation

$$\hat{\beta}_P = \hat{\beta}_{AU} f_1^{KS-2} \left(\hat{\beta}_P (m_{(N)} - m_{\min}) \right).$$

Using the estimator $\bar{m}_{(n)}$ for $E(M_{(n)} | \infty)$, where $1 \leq n \leq N$, we have a generalized Aki-Utsu estimator

$$\hat{\beta}_{GAU}^{(n)} = \frac{H_n}{\bar{m}_{(n)} - m_{\min}}.$$

In the case $n=1$ holds $\hat{\beta}_{AU} = \hat{\beta}_{GAU}^{(1)}$.

The Page' estimator

Similar way like the generalized Aki-Utsu estimator the Page' estimator can be solved by using

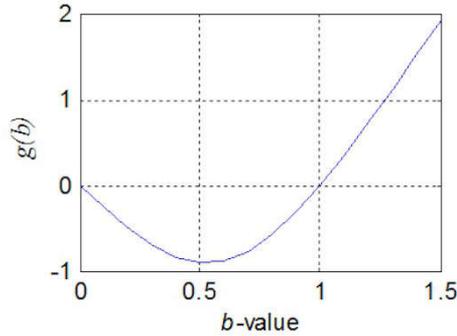
$$g_1 \left(\tilde{\beta}_{GP}^{(n)} \right) = -\tilde{\beta}_{GP}^{(n)} (m_{\max} - \bar{m}_{(n)}) + f_n^{KS-1} \left(\tilde{\beta}_{GP}^{(n)} (m_{\max} - m_{\min}) \right)$$

or

$$g_2 \left(\tilde{\beta}_{GP}^{(n)} \right) = \tilde{\beta}_{GP}^{(n)} (\bar{m}_{(n)} - m_{\min}) - f_n^{KS-2} \left(\tilde{\beta}_{GP}^{(n)} (m_{\max} - m_{\min}) \right).$$

Actually these functions are exactly the same since $g_2 \left(\tilde{\beta}_{GP}^{(n)} \right) - g_1 \left(\tilde{\beta}_{GP}^{(n)} \right) = 0$. We call this a generalized Page' estimator $\tilde{\beta}_{GP}^{(n)}$. In the case $n=1$ holds $\tilde{\beta}_P = \tilde{\beta}_{GP}^{(1)}$.

In Figure 1 we show an example curve of an auxiliary function g with parameters $b=1$, $m_{\max}=8$, $m_{\min}=5$ and $n=40$. The function g has a parabolic form, so it can be solved with Ridder's or Newton-Raphson method (Press et al., 1992).


 Figure 1: Example of the auxiliary function g

When making simulations (Haarala and Orosco, 2016) we showed how numerous artificial catalogues failed because of the condition $E(M_{(n)} | m_{\max}) \leq m_{\min} + H_n/\beta$. When we handle real seismic catalogues it is first calculated the estimator $\hat{\beta}$ and then with this estimator it is evaluated the estimator for the maximum. Due to the KS-2 function gives unique solution of the $\hat{\beta}^{(n)}$ -value for each maximum m_{\max} , it can be said that $\hat{\beta}^{(n)}$ is a function of m_{\max} (see Appendix A)

$$\hat{\beta}^{(n)}(m_{\max})(\bar{m}_{(n)} - m_{\min}) = f_n^{KS-2}(\hat{\beta}^{(n)}(m_{\max})(m_{\max} - m_{\min})).$$

So the Page' estimator is $\hat{\beta}_p = \hat{\beta}^{(1)}(m_{(N)})$ (where $m_{(N)}$ is the maximum observed value of catalogue) and the Aki-Utsu estimator $\hat{\beta}_{AU} = \hat{\beta}^{(1)}(\infty)$. These estimators hold the inequality

$$\hat{\beta}_{GP}^{(n)} = \hat{\beta}^{(n)}(m_{(n)}) \leq \hat{\beta}^{(n)}(\hat{m}_{\max}) \leq \hat{\beta}^{(n)}(\infty) = \hat{\beta}_{GAU}^{(n)}.$$

This shows that the β -value measures the distance between the minimum and the infinity similar way than the expected value of maximum

$$E(M_{(n)} | m_{(n)}) \leq E(M_{(n)} | \hat{m}_{\max}) \leq E(M_{(n)} | \infty).$$

It is well known that when $\beta \rightarrow \hat{\beta}_{AU}$ then $m_{\max} \rightarrow \infty$. (Normally this is presented in an opposite way as $m_{\max} \rightarrow \infty$ then $\hat{\beta}_{AU} \rightarrow \beta$.) A lesser known result is a limit when $\beta \rightarrow 0$. Let consider the KS-2 function

$$E(M_{(n)} | m_{\max}) - m_{\min} = \frac{n}{\beta} \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^k}{k(k+n)}. \quad (16)$$

Because of

$$1 - \exp[-\beta(m_{\max} - m_{\min})] = \beta(m_{\max} - m_{\min}) \left\{ 1 + \sum_{j=2}^{\infty} \frac{[-\beta(m_{\max} - m_{\min})]^{j-1}}{j!} \right\}$$

the equation (16) results

$$E(M_{(n)} | m_{\max}) - m_{\min} = \frac{n}{n+1} (m_{\max} - m_{\min}) \left\{ 1 + \sum_{j=2}^{\infty} \frac{[-\beta(m_{\max} - m_{\min})]^{j-1}}{j!} \right\} \\ + n \sum_{k=2}^{\infty} \frac{\beta^{k-1} \left((m_{\max} - m_{\min}) \left\{ 1 + \sum_{j=2}^{\infty} \frac{[-\beta(m_{\max} - m_{\min})]^{j-1}}{j!} \right\} \right)^k}{k(k+n)}.$$

If $\beta \rightarrow 0$ then we have

$$m_{\max} = m_{\min} + \frac{n+1}{n} [E(M_{(n)} | m_{\max}) - m_{\min}]. \quad (17)$$

For example, in the case $n = 2$ this gives $m_{\max} \geq m_{(1)} + 3/2(m_{(2)} - m_{(1)}) = m_{(2)} + (m_{(2)} - m_{(1)})/2$. Since it is a limit as $\beta \rightarrow 0$ it does not mean that the maximum is just two times the distance between minimum and mean. Actually is a minimum distance to the maximum i.e.

$$m_{\max} \geq m_{\min} + \frac{n+1}{n} (\bar{m}_{(n)} - m_{\min})$$

Up to this point we can apply these concepts in the case we have a seismic catalogue where we take the two biggest magnitudes, say for example 8.8 and 9.5, so this is our sub-catalogue of size $n = 2$. We have then $\hat{m}_{\max} \geq 9.5 + (9.5 - 8.8)/2 = 9.85$. This is a statistical estimator so even though the formula holds that the maximum is equal or bigger than the result it gives, it could be smaller. The advantage of this approach is that we do not need to know the β -value to have some estimation of the maximum and besides, it is easy to calculate.

We could also find the minimum convergence magnitude to the Page' estimator. The equation(17) states now that there is solution only if

$$\bar{m}_{(n)} < m_{\min} + \frac{n}{n+1} (m_{\max} - m_{\min}). \quad (18)$$

where m_{\max} can be replaced with its estimator.

Actually there is no numerical solution when the equality holds i.e. when $\beta = 0$. That is to say that if the estimator of the expected value $\bar{m}_{(n)}$ is equal or bigger than this limit then the estimator $\hat{\beta}_{GP}^{(n)}$ has no solution. In that case the $\hat{\beta}$ value will be zero or negative when the inequality fails so we can define $\hat{\beta}_{GP}^{(n)} = 0$ in those cases.

In figure 2 we plotted the Page' estimator $\hat{\beta}_{GP}^{(n)}$ and the Aki-Utsu estimator $\hat{\beta}_{GAU}^{(n)}$ for the case: $b = 1, m_{\max} = 8, m_{\min} = 5$ and $n = 200$. There is only one artificial catalogue, but the estimators calculated considering $1 < n < N$ have been «smoothed» taking 1000 random sets of sub-catalogues and calculating the mean for them. We can see that basically $[\hat{\beta}_{GP}^{(n)}, \hat{\beta}_{GAU}^{(n)}]$ grows when the size of sub-catalogue grows. In the very first sub-catalogues the correct β is not included into the interval. The figure 2 has been made with Ridder's method (Press et al, 1992) starting with the interval $[\log(10)/10^8, \hat{\beta}_{GAU}^{(n)} + \log(10)/10]$.

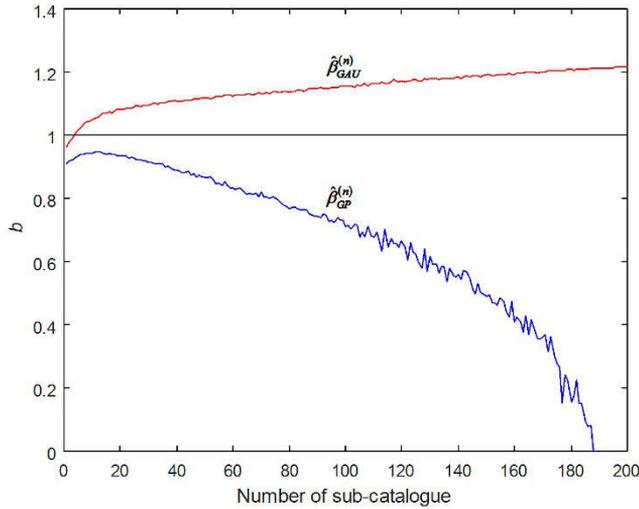


Figure 2: Page' and Aki-Utsu estimators with different sub-catalogue sizes using Ridders' method.

The exact variance

Comparing to the calculus of the first moment (Haarala and Orosco, 2016), the calculus of the variance is much more complicated. Integrating by parts it is

$$E\left(M_{(n)}^2 \mid m_{\max}\right) = \int_{m_{\min}}^{m_{\max}} m^2 dF_{M_{(n)}}(m \mid m_{\max}) = m_{\max}^2 - 2 \int_{m_{\min}}^{m_{\max}} m F_{M_{(n)}}(m \mid m_{\max}) dm. \quad (19)$$

When we applied the functions (1) and (2) the integral can be written

$$\int_{m_{\min}}^{m_{\max}} m F_{M_{(n)}}(m \mid m_{\max}) dm = \frac{\int_{m_{\min}}^{m_{\max}} m \left(1 - \exp[-\beta(m - m_{\min})]\right)^n dm}{\left(1 - \exp[-\beta(m_{\max} - m_{\min})]\right)^n}.$$

Because of the integrate function of $\left(1 - \exp[-\beta(m - m_{\min})]\right)^n$ is

$$\int \left(1 - \exp[-\beta(m - m_{\min})]\right)^n dm = \left[1 - \exp[-\beta(m - m_{\min})]\right]^n \frac{1}{\beta} \int_n^{KS-1} (\beta(m - m_{\min})) \quad (20)$$

(the proof has been given in the Appendix B), integrating again by parts the second moment (19) yields to

$$E(M_{(n)}^2 | m_{max}) = m_{max}^2 - \frac{2m_{max}}{\beta} f_n^{KS-1}(\beta(m_{max} - m_{min})) + \frac{2}{\beta} \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} \right]^n f_n^{KS-1}(\beta(m - m_{min})) dm. \quad (21)$$

Using notation $\Delta_n(m) = \beta^{-1} f_n^{KS-1}(\beta(m - m_{min}))$ the second moment is

$$E(M_{(n)}^2 | m_{max}) = m_{max}^2 - 2m_{max} \Delta_n(m_{max}) + 2 \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} \right]^n \Delta_n(m) dm.$$

Hence

$$\begin{aligned} [E(M_{(n)} | m_{max})]^2 &= [m_{max} - \Delta_n(m_{max})]^2 \\ &= m_{max}^2 - 2m_{max} \Delta_n(m_{max}) + \Delta_n^2(m_{max}) \end{aligned}$$

so the variance can be written

$$\begin{aligned} Var(M_{(n)} | m_{max}) &= E(M_{(n)}^2 | m_{max}) - [E(M_{(n)} | m_{max})]^2 \\ &= 2 \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} \right]^n \Delta_n(m) dm - \Delta_n^2(m_{max}). \end{aligned} \quad (22)$$

Since for all m holds $\Delta_n(m) \leq \Delta_n(m_{max})$ the variance has an upper limit

$$Var(M_{(n)} | m_{max}) \leq 2\Delta_n(m_{max}) \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} \right]^n dm - \Delta_n^2(m_{max}) = \Delta_n^2(m_{max}).$$

Next we need to solve the integral in the equation(21). Assuming that n is integer the KS-1 function can be written using finite sum as (see Haarala and Orosco, 2016)

$$f_n^{KS-1}(\beta(m - m_{min})) = \frac{\beta(m - m_{min}) - \sum_{k=1}^n (1 - \exp[-\beta(m - m_{min})])^k}{(1 - \exp[-\beta(m - m_{min})])^n}.$$

Now the integral in (21) can be written

$$\frac{2}{\beta} \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} \right]^n f_n^{KS-1}(\beta(m - m_{min})) dm =$$

$$\begin{aligned}
 & \frac{2}{\beta} \int_{m_{\min}}^{m_{\max}} \beta(m - m_{\min}) - \sum_{k=1}^n \frac{(1 - \exp[-\beta(m - m_{\min})])^k}{k} dm \\
 &= \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^n}{\beta(m_{\max} - m_{\min})^2 - \frac{2}{\beta} \sum_{k=1}^n \left\{ \frac{1}{k} \int_{m_{\min}}^{m_{\max}} (1 - \exp[-\beta(m - m_{\min})])^k dm \right\}} \\
 &= \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^n}{\beta(m_{\max} - m_{\min})^2 - \frac{2}{\beta} \sum_{k=1}^n \left\{ \frac{1}{k} \int_{m_{\min}}^{m_{\max}} (1 - \exp[-\beta(m - m_{\min})])^k dm \right\}}
 \end{aligned}$$

The integral in the sum is given in (20). We can write the exact variance as

$$\begin{aligned}
 \text{Var}(M_{(n)} | m_{\max}) &= \frac{\beta^2 (m_{\max} - m_{\min})^2 - 2 \sum_{k=1}^n \left\{ \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k} \int_k^{KS-1} (\beta(m_{\max} - m_{\min})) \right\}}{\beta^2 (1 - \exp[-\beta(m_{\max} - m_{\min})])^n} - \Delta_n^2(m_{\max}) \quad (23) \\
 &= \frac{\Delta_0^2(m_{\max}) - \frac{2}{\beta} \sum_{k=1}^n \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k \Delta_k(m_{\max})}{k}}{(1 - \exp[-\beta(m_{\max} - m_{\min})])^n} - \Delta_n^2(m_{\max}).
 \end{aligned}$$

Without applying series this variance looks like (the proof is given in the Appendix C)

$$\begin{aligned}
 \text{Var}(M_{(n)} | m_{\max}) &= \frac{[\beta(m_{\max} - m_{\min}) - H_n]^2 - H_n^2 + 2 \sum_{j=0}^{n-1} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k(k+j)}}{\beta^2 (1 - \exp[-\beta(m_{\max} - m_{\min})])^n} \\
 &\quad - \left\{ \frac{\beta(m_{\max} - m_{\min}) - \sum_{k=1}^n \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k}}{\beta [1 - \exp[-\beta(m_{\max} - m_{\min})]]^n} \right\}^2. \quad (24)
 \end{aligned}$$

Both expressions of the variance (equations(23) and (24)) are numerically unstable. The numerically stable form of variance can be written as (proof in Appendix D)

$$\text{Var}(M_{(n)} | m_{\max}) = \frac{1}{\beta^2} \sum_{k=2}^{\infty} \frac{2n}{2n+k} \left\{ \sum_{j=1}^{k-1} \frac{1}{n+j} \right\} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^k}{n+k}. \quad (25)$$

The series can be solved similar way we made with the KS-1 and KS-2 (Haarala and Orosco, 2016). Moreover, the expression (25) has continuous variable n . That means that we could replace n with $T\lambda$, where T is a continuous variable of time and λ is a rate of events in some time unit (normally it is a year).

The variance(25) can be calculated by means of our MATLAB function (Haarala and Orosco, 2016) as

$$\text{Var}(M_{(n)} | m_{\max}) = \frac{1}{\beta^2} \text{KS}(3, \beta(m_{\max} - m_{\min}), n).$$

In Figure 3 we show an example of the calculus of the variance for the case when $b = 1$, $m_{\max} = 8$ and $m_{\min} = 5$ given different sizes of catalogues. The variance has the shape of the Gamma function. The maximum value is at $n = 7$; from size $n = 65$ on, the level of variance is less than the level of variance when the catalogue size is $n = 1$.

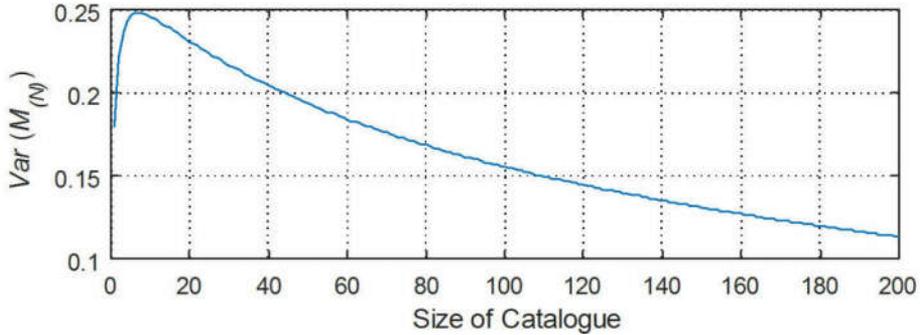


Figure 3: Example curve of the variance.

More about the variance

As we have shown, the first moment is related with the maximum likelihood estimators (Aki, 1965; Page, 1968) and other moment estimators (Hamilton, 1967; Cosentino et al, 1977), so it is quite expected that the exact variance will be the same in the case $n = 1$.

In the formula (24) we set $n = 1$, then we have

$$\text{Var}(M_{(1)} | m_{\max}) = \frac{1}{\beta^2} \left[\frac{m_{\max} - m_{\min}}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \right]^2 \exp[-\beta(m_{\max} - m_{\min})] \quad (26)$$

We can find the variance given by Aki (1965) as $m_{\max} \rightarrow \infty$:

$$\text{Var}(M_{(1)} | \infty) = \frac{1}{\beta^2}.$$

Generalizing, the variance at $m_{\max} = \infty$ is

$$\text{Var}(M_{(n)} | \infty) = \frac{1}{\beta^2} \sum_{k=1}^n \frac{1}{k^2}.$$

We can get it starting from (25). The proof is given in the Appendix E. It shows that the variance

is a bounded function for all m_{\max} (and also for all n) and it is smaller than $1.65 / \beta^2$.

The variance (26) is equal to the variance as it was given by Hamilton (1967)

$$Var(M_{(1)} | m_{\max}) = \frac{1}{\beta^2} \left[\frac{m_{\max} - m_{\min}}{\exp\left[\frac{\beta(m_{\max} - m_{\min})}{2}\right] - \exp\left[-\frac{\beta(m_{\max} - m_{\min})}{2}\right]} \right]^2$$

(actually Hamilton wrote $\exp[z] - \exp[-z] = 2 \sinh[z]$) and the variance was given by Cosentino et al (1976, 1977)

$$\begin{aligned} Var(M_{(1)} | m_{\max}) &= \frac{1}{\beta^2} \left[m_{\max} + \frac{1}{\beta} - \bar{m} \right]^2 \exp[-\beta(m_{\max} - m_{\min})] \quad (\text{ver. 1976}) \\ &= \frac{1}{\beta^2} \left[m_{\min} + \frac{1}{\beta} - \bar{m} \right]^2 \exp[\beta(m_{\max} - m_{\min})] \quad (\text{ver. 1977}) \end{aligned}$$

where they used the equation (10) to replace the square of brackets in (26).

We have showed now how variance fits with the models of Hamilton (1967) and Cosentino et al (1976; 1977). It is natural since the distribution function is the same in the case $n = 1$. If we write the variance (26) as

$$Var(M_{(1)} | m_{\max}) = \frac{\partial}{\partial \beta} \left\{ -\frac{1}{\beta} + \frac{(m_{\max} - m_{\min})}{1 - \exp[-\beta(m_{\max} - m_{\min})]} + C \right\}$$

where C is some constant, for example $C = \bar{m} - m_{\max}$. Using the equation of the variance of Cosentino et al (1976) (10), we have some function g such that

$$g(\beta) = \bar{m} - m_{\max} - \frac{1}{\beta} + \frac{m_{\max} - m_{\min}}{1 - \exp[-\beta(m_{\max} - m_{\min})]}.$$

Now $g(\beta) = 0$ because of the equation (10), and $g'(\beta) = Var(M_{(1)} | m_{\max})$. So the g is maximum likelihood function L

$$\begin{aligned} g(\beta) &= \frac{1}{N} \frac{\partial}{\partial \beta} L(\beta; m) \\ &= \frac{1}{N} \frac{\partial}{\partial \beta} \left\{ \log \prod_{k=1}^N \frac{\beta \exp[-\beta(m_k - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \right\}. \end{aligned} \quad (27)$$

Simulations

In figure 4 we display the mean estimator for the expected value of $M_{(n)}$ given by (6). We generated only one random catalogue C_{200} with parameters $b = 1$, $m_{\max} = 8$, $m_{\min} = 5$ and $n = 200$ (figure 4a); using this catalogue we take randomly the events from the catalogue C_{200} into the sub-catalogues $C_{n,k}$. The means $\bar{m}_{(1)}$ and $\bar{m}_{(200)}$ are fixed. For the sub-catalogue of sizes $1 < n < 200$ we get different paths. The continuous thick line represents the mean value of 1000 samples and the dashed line is the theoretical expected value. We can see that the mean values are quite close to the expected values in the range of small catalogue sizes. Actually this is quite expected since for example for the n -values 1, 2 and 3 the number of sub-catalogues are 200, 100 and 66, respectively, and the estimator of the expected value is a mean value of the maximums of those sub-catalogues. The figure 4b is the same than the 4a but it is generated with parameters $b = 1$, $m_{\max} = 8$, $m_{\min} = 6$ and $n = 200$.

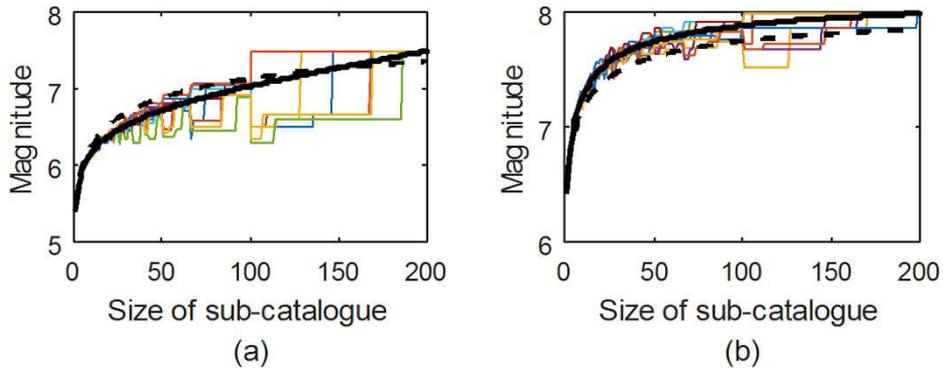


Figure 4: The estimator(6) for some catalogue C_{200} with different parameters

We generated 10000 catalogues with parameters $b = 1$, $m_{\max} = 8$, $m_{\min} = 5$ and $n = 200$. For those catalogues we have calculated the mean $\bar{m}_{(n)}$. In the figures 5a, 5b, 5c and 5d we show the histograms of means $\bar{m}_{(1)}$, $\bar{m}_{(10)}$, $\bar{m}_{(100)}$ and $\bar{m}_{(200)}$ respectively.

Also in the figures is pointed out a limit $E(M_{(n)} | \infty)$ when we set $b = 1$ (white line). We can see how close this limit is to the expected value, which is estimated with the mean value. We can also see that the distribution function is

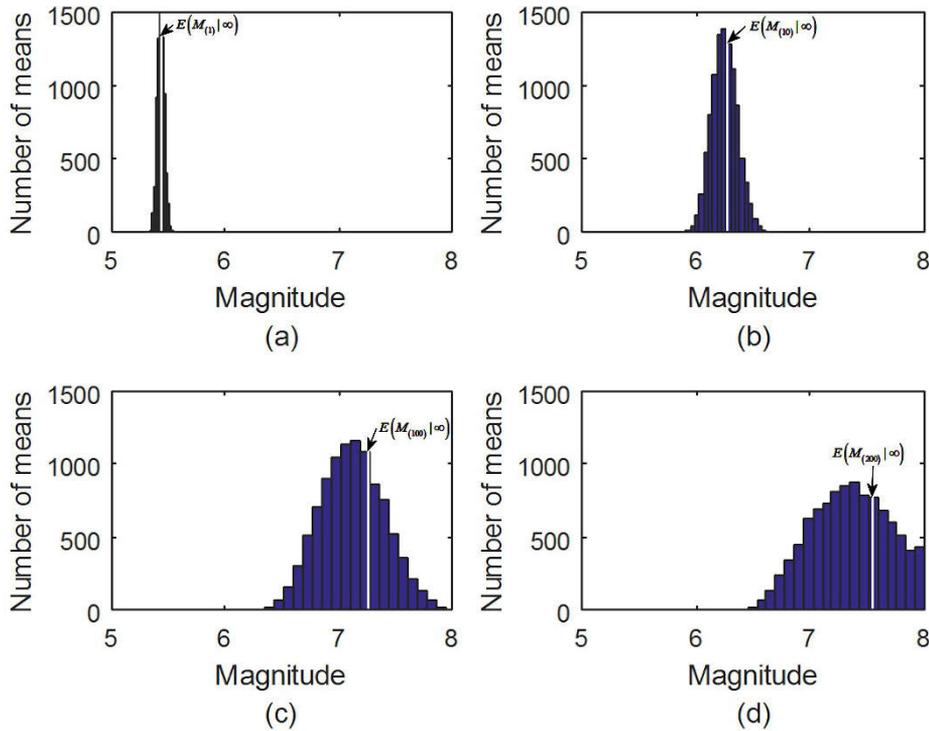


Figure 5: Distribution of mean values

normal (as it should be because of the central limit theorem) except in the figure 5d. Because the probability that the event is outside of interval $[m_{\min}, m_{\max}]$ is zero, the probability distribution function is actually a double truncated normal distribution function. Of course the normalization factor is about one for all cases except in the case 5d, so we can use the unbounded normal distribution function in those cases. Moreover, we can see that the variance is much smaller in case 5a than cases 5b-5d.

In the figure 6a we plotted a mean of the estimators \hat{m}_{\max} using the same β -value (as we made in the Part I) that we used to generate the artificial catalogues, and the Page' estimator $\hat{\beta}_p$ (i.e. Page' estimator is calculated at $n=1$ using maximum observed value $m_{(N)}$). We have got Figure 6a by generating 1000 artificial catalogues with parameters $b=1$, $m_{\max}=8$, $m_{\min}=5$ and $1 \leq n \leq 200$. The Figure 6b shows how many catalogues could be used to calculate these mean values.

Because of the Page' estimator underestimates the β -value we can use more catalogues in the calculus but still 25-30% catalogues are rejected, which provokes the bias to the mean value of the maximums.

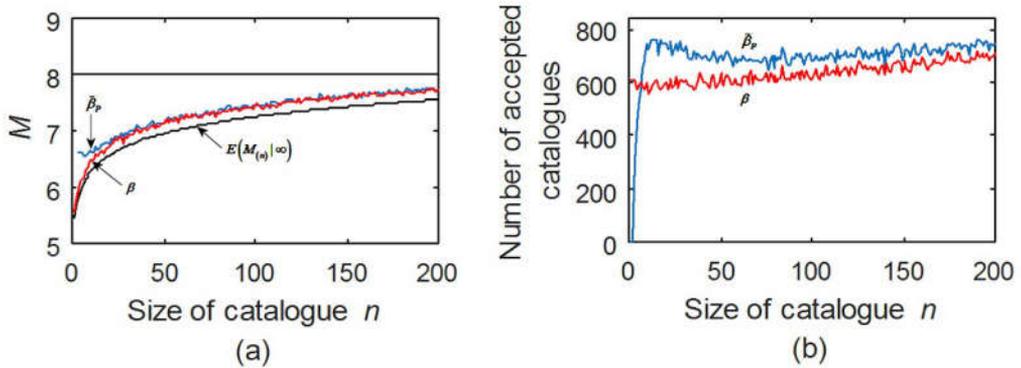


Figure 6: Estimation of the maximum \hat{m}_{\max}

Concluding remarks

In this work we showed that the Aki and Page's (Maximum Likelihood) estimators and Hamilton, Cosentino et al. and Utsu's (Moment) estimators, can be evaluated by using KS functions. When the sub-catalogue size is $n = 1$, all the estimators gave the same values when we use the same parameters. The differences, if any, come from the computation.

In order statistic we can generalize these results to the family of estimators. Due to the method we are here proposing joins with moment estimator method when $n = 1$, we could solve the estimators β and m_{\max} using first and second moments, even though we do not use this approach in this series of works, but it was made in earlier ones. Any way we showed that the order statistic carries more information than Moment Estimator Method, because we could apply the Moment Estimator Method for each $1 \leq n \leq N$.

Appendix A

We consider the function

$$y = f_n^{KS-2}(x) = n \sum_{k=1}^{\infty} \frac{(1 - \exp[-x])^k}{k(k+n)}. \quad (28)$$

It has a derivative

$$\begin{aligned} (f_n^{KS-2})'(x) &= n \exp[-x] \sum_{k=1}^{\infty} \frac{(1 - \exp[-x])^{k-1}}{(k+n)} \\ &= n \exp[-x] \left[\frac{1}{(1+n)} + f_{n+1}^{KS-1}(x) \right]. \end{aligned}$$

We can see that $(f_n^{KS-2})'(x) > 0$ for all $x \in [0, \infty[$, so the function $f_n^{KS-2}(x)$ is a strictly increasing function. This means that it is a bijection from $[0, \infty[$ onto $[0, H_n[$ and it has a unique solution.

Let's consider now the equation

$$\beta(\bar{m}_{(n)} - m_{\min}) = f_n^{KS-2}(\beta(m_{\max} - m_{\min})). \quad (29)$$

Assume that there are some $\hat{\beta}_1^{(n)}, \hat{\beta}_2^{(n)}$ and $m_{\max,1}, m_{\max,2}$, respectively. If

$$f_n^{KS-2}(\hat{\beta}_1^{(n)}(m_{\max,1} - m_{\min})) = f_n^{KS-2}(\hat{\beta}_2^{(n)}(m_{\max,2} - m_{\min}))$$

then $\hat{\beta}_1^{(n)}(\bar{m}_{(n)} - m_{\min}) = \hat{\beta}_2^{(n)}(\bar{m}_{(n)} - m_{\min})$ and $\hat{\beta}_1^{(n)} = \hat{\beta}_2^{(n)}$. Moreover, since $f_n^{KS-2}(x)$ is the bijection, then $\hat{\beta}_1^{(n)}(m_{\max,1} - m_{\min}) = \hat{\beta}_2^{(n)}(m_{\max,2} - m_{\min})$. But now $\hat{\beta}_1^{(n)} = \hat{\beta}_2^{(n)}$, so also $m_{\max,1} = m_{\max,2}$. That is to say, if $\hat{\beta}_1^{(n)}, \hat{\beta}_2^{(n)}$ and $m_{\max,1}, m_{\max,2}$ are from $y = f_n^{KS-2}(x)$ then $m_{\max,1} = m_{\max,2}$ and $\hat{\beta}_1^{(n)} = \hat{\beta}_2^{(n)}$. This makes the one-to-one mapping between $\hat{\beta}^{(n)}$ and m_{\max} .

For each $\beta = \hat{\beta}^{(n)} \in]0, \hat{\beta}_{AU}^{(n)}[$ there is an image $m_{\max} \in]m_{\min} + (n+1)(\bar{m}_{(n)} - m_{\min})/n, \infty[$ and for each we can find. This follows the reality that $\beta \rightarrow \hat{\beta}_{AU}^{(n)}$ when $m_{\max} \rightarrow \infty$, and $m_{\max} \rightarrow m_{\min} + (n+1)(\bar{m}_{(n)} - m_{\min})/n$ when $\beta \rightarrow 0$. Thus the equation (29) defines the bijection mapping between the sets $]0, \hat{\beta}_{AU}^{(n)}[$ and $]m_{\min} + (n+1)(\bar{m}_{(n)} - m_{\min})/n, \infty[$.

Appendix B

Firstly, it is to note that the derivative of the KS-2 function is

$$\begin{aligned} \frac{\partial}{\partial \mathfrak{M}} f_n^{KS-2}(\beta(\mathfrak{M} - m_{\min})) &= n \sum_{k=1}^{\infty} \frac{\partial}{\partial \mathfrak{M}} \frac{(1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^k}{k(k+n)} \\ &= n\beta \exp[-\beta(\mathfrak{M} - m_{\min})] \sum_{k=1}^{\infty} \frac{k(1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^{k-1}}{k(k+n)} \\ &= n\beta \frac{\exp[-\beta(\mathfrak{M} - m_{\min})]}{1 - \exp[-\beta(\mathfrak{M} - m_{\min})]} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})). \end{aligned}$$

Thus the derivative of the KS-1 function is

$$\begin{aligned} \frac{1}{\beta} \frac{\partial}{\partial \mathfrak{M}} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})) &= \frac{\partial}{\partial \mathfrak{M}} \left[(\mathfrak{M} - m_{\min}) - \frac{1}{\beta} f_n^{KS-2}(\beta(\mathfrak{M} - m_{\min})) \right] \\ &= 1 - n \frac{\exp[-\beta(\mathfrak{M} - m_{\min})]}{1 - \exp[-\beta(\mathfrak{M} - m_{\min})]} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})). \end{aligned}$$

So the direct derivative gives

$$\begin{aligned} \frac{\partial}{\partial \mathfrak{M}} \left[(1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^n \frac{1}{\beta} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})) \right] \\ &= n \exp[-\beta(\mathfrak{M} - m_{\min})] (1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^{n-1} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})) \\ &\quad + (1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^n \left\{ 1 - n \frac{\exp(-\beta(\mathfrak{M} - m_{\min}))}{1 - \exp(-\beta(\mathfrak{M} - m_{\min}))} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})) \right\} \\ &= (1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^n, \end{aligned}$$

and we have

$$\int (1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^n d\mathfrak{M} = (1 - \exp[-\beta(\mathfrak{M} - m_{\min})])^n \frac{1}{\beta} f_n^{KS-1}(\beta(\mathfrak{M} - m_{\min})) + C. \quad (30)$$

Appendix C

Because of

$$\begin{aligned} [1 - \exp[-\beta(\varpi - m_{\min})]]^k &= [1 - \exp[-\beta(\varpi - m_{\min})]]^{k-1} \\ &\quad - [1 - \exp[-\beta(\varpi - m_{\min})]]^{k-1} \exp[-\beta(\varpi - m_{\min})] \end{aligned}$$

so the integral

$$\begin{aligned} &\int_{m_{\min}}^{m_{\max}} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(\varpi - m_{\min})]]^k}{k+j} d\varpi \\ &= \int_{m_{\min}}^{m_{\max}} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(\varpi - m_{\min})]]^{k-1}}{k+j} d\varpi - \frac{1}{\beta} \sum_{k=1}^{n-j} \int_{m_{\min}}^{m_{\max}} \frac{[1 - \exp[-\beta(\varpi - m_{\min})]]^{k-1} \beta \exp[-\beta(\varpi - m_{\min})]}{k+j} d\varpi \\ &= \frac{m_{\max} - m_{\min}}{j+1} + \int_{m_{\min}}^{m_{\max}} \sum_{k=1}^{n-(j+1)} \frac{[1 - \exp[-\beta(\varpi - m_{\min})]]^k}{k+(j+1)} d\varpi - \frac{1}{\beta} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k(k+j)}, \end{aligned}$$

where $j = 0, 1, \dots, n-1$, is applied n times; then it gives for the integral

$$\begin{aligned} &\frac{2}{\beta} \int_{m_{\min}}^{m_{\max}} \sum_{k=1}^n \frac{[1 - \exp[-\beta(\varpi - m_{\min})]]^k}{k} d\varpi \\ &= \frac{2}{\beta} (m_{\max} - m_{\min}) H_n - \frac{2}{\beta^2} \sum_{j=0}^{n-1} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k(k+j)}. \end{aligned}$$

Now the variance gets the form

$$\begin{aligned} \text{Var}(M_{(n)} | m_{\max}) &= \frac{\beta^2 (m_{\max} - m_{\min})^2 - 2\beta(m_{\max} - m_{\min})H_n + 2 \sum_{j=0}^{n-1} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k(k+j)}}{\beta^2 (1 - \exp[-\beta(m_{\max} - m_{\min})])^n} - \Delta_n^2(m_{\max}) \\ &= \frac{[\beta(m_{\max} - m_{\min}) - H_n]^2 + 2 \sum_{j=0}^{n-1} \sum_{k=1}^{n-j} \frac{[1 - \exp[-\beta(m_{\max} - m_{\min})]]^k}{k(k+j)} - H_n^2}{\beta^2 (1 - \exp[-\beta(m_{\max} - m_{\min})])^n} - \Delta_n^2(m_{\max}). \end{aligned}$$

Appendix D

To get the numerically stable form for the variance, we start from (22)

$$\text{Var}(M_{(n)} | m_{\max}) = \frac{2}{\beta} \int_{m_{\min}}^{m_{\max}} \left[\frac{1 - \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \right]^n f_n^{KS-1}(\beta(m - m_{\min})) dm - \left[\frac{1}{\beta} f_n^{KS-1}(\beta(m_{\max} - m_{\min})) \right]^2.$$

The integral can be written as

$$\begin{aligned} & \frac{2}{\beta} \int_{m_{\min}}^{m_{\max}} \left[\frac{1 - \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \right]^n f_n^{KS-1}(\beta(m - m_{\min})) dm \\ &= \frac{2}{\beta} \int_{m_{\min}}^{m_{\max}} \beta(m - m_{\min}) - \sum_{k=1}^n \frac{(1 - \exp[-\beta(m - m_{\min})])^k}{k} dm \\ &= \frac{2}{\beta} \sum_{k=n+1}^{\infty} \left\{ \frac{1}{k} \int_{m_{\min}}^{m_{\max}} (1 - \exp[-\beta(m - m_{\min})])^k dm \right\} \\ &= \frac{2}{\beta} \sum_{k=n+1}^{\infty} \left\{ \frac{1}{k} (1 - \exp[-\beta(m_{\max} - m_{\min})])^{k-n} \frac{1}{\beta} f_k^{KS-1}(\beta(m_{\max} - m_{\min})) \right\} \\ &= \frac{2}{\beta^2} \sum_{k=1}^{\infty} \left\{ \frac{1}{k+n} (1 - \exp[-\beta(m_{\max} - m_{\min})])^k f_{k+n}^{KS-1}(\beta(m_{\max} - m_{\min})) \right\} \end{aligned} \quad (31)$$

because

$$\beta(m - m_{\min}) = \sum_{k=1}^{\infty} \frac{(1 - \exp[-\beta(m - m_{\min})])^k}{k} = f_0^{KS-1}(\beta(m_{\max} - m_{\min}))$$

and using the integral formula(30) we get

$$f_m^{KS-1}(\beta(m_{\max} - m_{\min})) = \sum_{j=1}^{\infty} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^j}{j + m},$$

then the last series in the(31) can be written as

$$\begin{aligned} \frac{2}{\beta^2} \sum_{k=1}^{\infty} \left\{ \frac{1}{k+n} (1 - \exp[-\beta(m_{\max} - m_{\min})])^k f_{k+n}^{KS-1}(\beta(m_{\max} - m_{\min})) \right\} \\ = \frac{2}{\beta^2} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^{k+j}}{(k+n)(j+k+n)}. \end{aligned} \tag{32}$$

Now it is to collect terms containing the same power so we have for some $k' = k + j$, $k' = 2, 3, 4, \dots$,

$$\begin{aligned} \sum_{j=1}^{k'-1} \frac{(1 - \exp[-\beta(m - m_{\min})])^{k'}}{(k' - j + n)(k' + n)} &= \frac{(1 - \exp[-\beta(m - m_{\min})])^{k'}}{k' + n} \sum_{j=1}^{k'-1} \frac{1}{n + k' - j} \\ &= \frac{(1 - \exp[-\beta(m - m_{\min})])^{k'}}{k' + n} \sum_{j'=1}^{k'-1} \frac{1}{n + j'}. \end{aligned}$$

Applying this into (32) and rewriting the index of series we get

$$\frac{2}{\beta^2} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \frac{(1 - \exp[-\beta(m - m_{\min})])^{k+j}}{(k+n)(j+k+n)} = \frac{2}{\beta^2} \sum_{k=2}^{\infty} \left\{ \sum_{j=1}^{k-1} \frac{1}{n+j} \right\} \frac{(1 - \exp[-\beta(m - m_{\min})])^k}{n+k}, \tag{33}$$

so the variance is given by

$$\begin{aligned} Var(M_{(n)} | m_{\max}) &= \frac{2}{\beta^2} \sum_{k=2}^{\infty} \left\{ \sum_{j=1}^{k-1} \frac{1}{n+j} \right\} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^k}{n+k} \\ &\quad - \left[\frac{1}{\beta} f_n^{KS-1}(\beta(m_{\max} - m_{\min})) \right]^2. \end{aligned} \tag{34}$$

We could also write $[\beta^{-1} f_n^{KS-1}(\beta(m_{\max} - m_{\min}))]^2$ similarly as the series(32)

$$\begin{aligned} \frac{1}{\beta^2} \sum_{k=1}^{\infty} \left\{ \frac{1}{k+n} (1 - \exp[-\beta(m_{\max} - m_{\min})])^k f_n^{KS-1}(\beta(m_{\max} - m_{\min})) \right\} \\ = \frac{1}{\beta^2} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^{k+j}}{(n+k)(n+j)}. \end{aligned} \tag{35}$$

In the same way as we got(33)above, the(35) gives

$$\left[\frac{1}{\beta} f_n^{KS-1}(\beta(m_{\max} - m_{\min})) \right]^2 = \frac{1}{\beta^2} \sum_{k=2}^{\infty} \left\{ \sum_{j=1}^{k-1} \frac{1}{(n+k-j)(n+j)} \right\} (1 - \exp[-\beta(m_{\max} - m_{\min})])^k. \tag{36}$$

Applying (36) into (34), we get for variance the formula

$$Var(M_{(n)} | m_{max}) = \frac{1}{\beta^2} \sum_{k=2}^{\infty} \left\{ \sum_{j=1}^{k-1} \frac{k+n-2j}{(n+j)(n+k-j)} \right\} \frac{(1 - \exp[-\beta(m_{max} - m_{min})])^k}{n+k}.$$

We could use this form of variance, but the inner sum produces a small error (bias) for the results because it is a sum of positive and negative values. To avoid this problem, we can calculate the pairs $(1, k-1), (2, k-2), \dots, (j, k-j)$. If k is even, then there is a $k/2$ term without pair. We have now

$$\begin{aligned} \frac{k+n-2j}{(n+j)(n+k-j)} + \frac{k+n-2(k-j)}{(n+(k-j))(n+k-(k-j))} &= \frac{2n}{(n+j)(n+k-j)} \\ \frac{k+n-2(k/2)}{(n+(k/2))(n+k-(k/2))} &= \frac{n}{(n+(k/2))(n+k-(k/2))} \end{aligned}$$

so we can write

$$\sum_{j=1}^{k-1} \frac{k+n-2j}{(n+j)(n+k-j)} = \sum_{j=1}^{k-1} \frac{n}{(n+j)(n+k-j)}. \quad (37)$$

The factor (37) is not also so good because we need to accelerate it. In the worst case the van Wijngaarden transformation uses the value $k = 2^{57}$ which means that the factor (37) would need a huge capacity of computation. The factor in the equation (34) is better, since it can accelerate its factor as

$$\sum_{j=1}^k \frac{1}{j+n} = \sum_{m=1}^{n+k} \frac{1}{m} - \sum_{m=1}^n \frac{1}{m} = H_{n+k} - H_n$$

where the harmonic number H_n can be calculated for example using Ramanujan's approximation (Villarino, 2008)

$$\begin{aligned} m &= \frac{n(n+1)}{2} \\ H_n &\approx \frac{1}{2} \log(2m) + \gamma + \frac{1}{12m} - \frac{1}{120m^2} + \frac{1}{630m^3} \\ &\quad - \frac{1}{1680m^4} + \frac{1}{2310m^5} - \frac{191}{360360m^6} \end{aligned} \quad (38)$$

for $n \geq 10$.

MATLAB has a *harmonic* function which bases to the Psi function (Appendix E). It is much slower than the Ramanujan's approximation. Only problem with the Ramanujan's approximation is that it is not so good for values less than 10. There is no sense to add more terms to the sum, because it grows the time of calculus. So when we have some value $n < 10$, $n \in \mathbb{R}$, then an easy trick is to find an integer D such that $n' = n + D \in [10, 11[$ when we have $H_n = H_{n'} - \sum_{j=1}^D 1/(n+j)$.

In the Figure 7 we show the absolutely error between the Ramanujan's approximation and

the Psi function calculating subtraction $d = H_n - H_{0.5}$, $n = 1.5, 2.5, 3.5, \dots, 9999.5$ with both method and calculating the difference of these results. Even the d is calculated with different methods, the difference between the results of the methods is less than 6×10^{-15} .

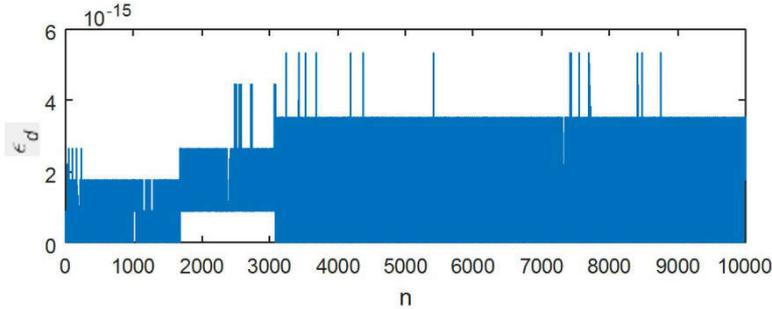


Figure 7: Error between Psi function method and Ramanujan's approximation method

To accelerate the factor (37) we need to modify it. It is clear that

$$\frac{1}{(n+j)(n+k-j)} = \frac{1}{2n+k} \left[\frac{1}{(n+j)} + \frac{1}{(n+k-j)} \right].$$

We get now

$$\begin{aligned} \sum_{j=1}^{k-1} \frac{n}{(n+j)(n+k-j)} &= \frac{n}{2n+k} \left[\sum_{j=1}^{k-1} \frac{1}{n+j} + \sum_{j=1}^{k-1} \frac{1}{n+k-j} \right] \\ &= \frac{2n}{2n+k} \sum_{j=1}^{k-1} \frac{1}{n+j}. \end{aligned}$$

The final variance can be written as

$$Var(M_{(n)} | m_{\max}) = \frac{1}{\beta^2} \sum_{k=2}^{\infty} \frac{2n}{2n+k} \left\{ \sum_{j=1}^{k-1} \frac{1}{n+j} \right\} \frac{(1 - \exp[-\beta(m_{\max} - m_{\min})])^k}{n+k}.$$

This is a positive term series (numerical stable) and we can accelerate it with the Ramanujan's approximation(38).

Appendix E

Let the variance to be at $m_{max} = \infty$

$$\beta^2 Var(M_{(n)} | m_{max} = \infty) = \sum_{k=2}^{\infty} \frac{2n}{(2n+k)(n+k)} \left\{ \sum_{j=1}^{k-1} \frac{1}{n+j} \right\}. \quad (39)$$

Because (39) is a series of nonnegative terms so it converges absolutely and then every rearrangement of the series converges (39), and they all converge to the same sum (Rudin, 1987). We see now that the terms of the series are

$$\begin{aligned} & \frac{2n}{(2n+2)(n+2)} \left\{ \frac{1}{n+1} \right\}, & k=2 \\ & \frac{2n}{(2n+3)(n+3)} \left\{ \frac{1}{n+1} + \frac{1}{n+2} \right\}, & k=3 \\ & \frac{2n}{(2n+4)(n+4)} \left\{ \frac{1}{n+1} + \frac{1}{n+2} + \frac{1}{n+3} \right\}, & k=4 \\ & \vdots \end{aligned}$$

so the rearrangement of the series(39) gives

$$\begin{aligned} \sum_{k=2}^{\infty} \frac{2n}{(2n+k)(n+k)} \left\{ \sum_{j=1}^{k-1} \frac{1}{n+j} \right\} &= \frac{2}{n+1} \sum_{k=1}^{\infty} \frac{n}{(2n+k+1)(n+k+1)} + \\ & \frac{2}{n+2} \sum_{k=1}^{\infty} \frac{n}{(2n+k+2)(n+k+2)} + \\ & \frac{2}{n+3} \sum_{k=1}^{\infty} \frac{n}{(2n+k+3)(n+k+3)} + \dots \\ &= \sum_{j=1}^{\infty} \frac{2}{n+j} \sum_{k=1}^{\infty} \frac{n}{(2n+k+j)(n+k+j)}. \end{aligned} \quad (40)$$

Owing to

$$\frac{1}{n+k+j} - \frac{1}{2n+k+j} = \frac{n}{(2n+k+j)(n+k+j)}$$

we can write

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{n}{(2n+k+j)(n+k+j)} &= \sum_{k=1}^{\infty} \frac{1}{n+k+j} - \sum_{k=1}^{\infty} \frac{1}{2n+k+j} \\ &= \sum_{k=1}^n \frac{1}{n+k+j} + \sum_{k=n+1}^{\infty} \frac{1}{n+k+j} - \sum_{k=1}^{\infty} \frac{1}{2n+k+j} \\ &= \sum_{k=1}^n \frac{1}{n+k+j} + \sum_{k=1}^{\infty} \frac{1}{2n+k+j} - \sum_{k=j+1}^{\infty} \frac{1}{2n+k+j} \\ &= \sum_{k=1}^n \frac{1}{n+k+j}. \end{aligned}$$

The series(40) gets now the form

$$\begin{aligned}
 \sum_{j=1}^{\infty} \frac{2}{n+j} \sum_{k=1}^{\infty} \frac{n}{(2n+k+j)(n+k+j)} &= \sum_{j=1}^{\infty} \frac{2}{n+j} \sum_{k=1}^n \frac{1}{n+j+k} \\
 &= 2 \sum_{k=1}^n \sum_{j=1}^{\infty} \frac{1}{(n+j)(n+j+k)} \\
 &= 2 \sum_{k=1}^n \left[\sum_{j=n+1}^{\infty} \frac{1}{j(j+k)} \right] \\
 &= 2 \sum_{k=1}^n \left[\frac{1}{k} \sum_{j=1}^{\infty} \frac{k}{j(j+k)} - \sum_{j=1}^n \frac{1}{j(j+k)} \right]
 \end{aligned} \tag{41}$$

Now the Psi function (Abramowitz and Stegun, 1972)

$$\begin{aligned}
 \psi(1+z) &= -\gamma + \sum_{j=1}^{\infty} \frac{z}{j(j+z)}, \quad z \neq -1, -2, -3, \dots, \\
 \psi(1) &= -\gamma, \quad \psi(n) = -\gamma + \sum_{j=1}^{n-1} \frac{1}{j}, \quad n \geq 2,
 \end{aligned}$$

where γ is the Euler-Mascheroni constant, gives

$$\sum_{j=1}^{\infty} \frac{k}{j(j+k)} = \gamma + \psi(1+k) = \sum_{j=1}^k \frac{1}{j}.$$

Applying this to (41) we get

$$\beta^2 \text{Var}(M_{(n)} | m_{\max} = \infty) = 2 \sum_{k=1}^n \left[\frac{1}{k} \sum_{j=1}^k \frac{1}{j} - \sum_{j=1}^n \frac{1}{j(k+j)} \right].$$

We will show now by induction that

$$2 \sum_{k=1}^n \left[\frac{1}{k} \sum_{j=1}^k \frac{1}{j} - \sum_{j=1}^n \frac{1}{j(k+j)} \right] = \sum_{k=1}^n \frac{1}{k^2}. \tag{42}$$

Let $n = 1$. Then we have

$$2 \sum_{k=1}^1 \left[\frac{1}{k} \sum_{j=1}^k \frac{1}{j} - \sum_{j=1}^n \frac{1}{j(k+j)} \right] = 1 = \sum_{k=1}^1 \frac{1}{k^2}.$$

Let assume now that (42) holds in the case n . Then at $n+1$ we have

$$\begin{aligned}
 2 \sum_{k=1}^{n+1} \left[\frac{1}{k} \sum_{j=1}^k \frac{1}{j} - \sum_{j=1}^{n+1} \frac{1}{j(k+j)} \right] &= 2 \sum_{k=1}^n \left[\frac{1}{k} \sum_{j=1}^k \frac{1}{j} - \sum_{j=1}^n \frac{1}{j(k+j)} \right] \\
 &\quad - 2 \sum_{k=1}^n \frac{1}{(n+1)(k+n+1)} + 2 \left[\sum_{j=1}^{n+1} \frac{1}{j(n+1)} - \sum_{j=1}^{n+1} \frac{1}{j(n+1+j)} \right] \\
 &= \sum_{k=1}^n \frac{1}{k^2} + 2 \left\{ \sum_{j=1}^{n+1} \frac{1}{(n+1)(n+1+j)} - \sum_{j=1}^n \frac{1}{(n+1)(n+1+j)} \right\} \\
 &= \sum_{k=1}^n \frac{1}{k^2} + \frac{1}{(n+1)^2} \\
 &= \sum_{k=1}^{n+1} \frac{1}{k^2}
 \end{aligned}$$

This completes the demonstration that

$$\text{Var}(M_{(n)} | \infty) = \frac{1}{\beta^2} \sum_{k=1}^n \frac{1}{k^2} = \frac{H_n^{(2)}}{\beta^2}$$

where $H_n^{(2)}$ is a harmonic number of order 2.

To the variance holds

$$\beta^2 \text{Var}(M_{(n)} | \infty) = \sum_{k=1}^n \frac{1}{k^2} = \zeta(2) = \frac{\pi^2}{6},$$

where $\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$ is a Riemann's zeta function (Abramowitz and Stegun, 1972). We see that for all $m_{max} \in [m_{min}, \infty]$ and $n = 1, 2, 3, \dots$ we will have

$$\text{Var}(M_{(n)} | m_{max}) \leq \frac{1}{\beta^2} \sum_{k=1}^n \frac{1}{k^2} < \frac{\pi^2}{6\beta^2} \approx \frac{1.6449}{\beta^2}.$$

so the variance is a bounded function.

References:

Aki, K. (1965). Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits, *Bull. Earthquake Res. Inst., Tokyo Univ.* **43**, 237-239.

Cao, A. M. and Gao, S. S. (2002). Temporal variation of seismic b -values beneath northeastern Japan island arc, *Geophys. Res. Lett.* **29** (9) 1-3.

Cosentino, P., and D. Luzio (1976). A generalization of the frequency-magnitude relation in the hypothesis of a maximum regional magnitude, *Ann. Geofis. (Rome)* **29**, 1-2, 3-8.

Cosentino, P., V. Ficara, and D. Luzio (1977). Truncated exponential frequency-magnitude

- relationship in the earthquake statistics, *Bull. Seism. Soc. Am.* **67**, 1615-1623.
- Haarala, M. and L. Orosco (2016). Analysis of Gutenberg-Richter b -value and m_{\max} . Part I: Exact solution of Kijko-Sellevoll estimator m_{\max} . *Cuaderno de Ingeniería. Nueva Serie.* [Salta, Argentina], núm 9, 2016: 51-67.
- Hamilton, R. M. (1967). Mean magnitude of an earthquake sequence, *Bull. Seism. Soc. Am.* **57**, 1115-1116.
- Kijko, A., and G. Graham (1998). "Parametric-historic" procedure for probabilistic seismic hazard analysis. Part I: Assessment of maximum regional magnitude m_{\max} , *Pure Appl. Geophys.* **152**, 413-442.
- Kijko, A. (2004), Estimation of the maximum earthquake magnitude, m_{\max} , *Pure Appl. Geophys.* **161**, 1-27.
- Page, R. (1968). Aftershocks and microaftershocks of the great Alaska earthquake of 1964, *Bull. Seism. Soc. Am.* **58**, 1131-1168.
- Pisarenko, V. F., A. A. Lyubushin, V. B. Lysenko, and T. V. Golubieva (1996). Statistical estimation of seismic hazard parameters: Maximum possible magnitude and related parameters, *Bull. Seism. Soc. Am.* **86**, 691-700.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical recipes in FORTRAN: The art of scientific computing*, 2nd ed., Cambridge University Press, Cambridge.
- Rudin, W. (1987). *Principles of Mathematical Analysis*, 3rd ed., 14th printing, McGraw-Hill, Singapore.
- Utsu, T. (1965). A method for determining the value of b in a formula $\log n = a - bM$ showing the magnitude-frequency relation for earthquakes, *Geophys. Bull. Hokkaido Univ.* **13**, 99-113.
- Villarino, M. B. (2008). Ramanujan's Harmonic Number Expansion into Negative Powers, *J. Inequal. Pure and Appl. Math.*, **9 (3)**, Art. 89, 12pp.

Diseño de un taller de introducción a la programación de robots¹

**M. Lorena Talamé - Norberto Aramayo - Ana E. Gardel -
M. Alicia Pérez Abelleira²**

Resumen

El interés por la robótica entre alumnos de la carrera de Ingeniería en Informática y la ausencia de temas directamente relacionados con la misma en el currículum, junto con la disponibilidad cada vez mayor de robots a costos accesibles, ha motivado el diseño de un taller de introducción a la programación de robots. Dicho taller permite a los alumnos incursionar en la disciplina usando tanto un robot real como un simulador, e integrar algunas habilidades ya adquiridas a lo largo de la carrera, con la motivación añadida de ver resultados al aplicarlas al control del robot. El presente trabajo describe las herramientas utilizadas y los contenidos, eminentemente prácticos del taller.

Palabras Clave: Robots móviles - Robótica educativa - Microsoft Robotics Development Studio - iRobot Create

Abstract

The increasing interest of informatics engineering students in learning about robotic and the lack of contents of this subject in currently courses, motivated the organization of the workshop. This activity let pupils to go inside the discipline, making use of a real robots and a simulator; at the same time, they can integrate knowledge acquired up to this time.

This report explains the tools used and the characteristics of the workshop, which has a primarily practical point of view.

Keywords: Mobil robots - educational robotics - Microsoft Robotics Development Studio - iRobot Create

¹ Este trabajo fue presentado en 2º Congreso Nacional de Ingeniería Informática/ Sistemas de Información, CoNaIISI 2014, San Luis, 13-14 Nov 2014. Publicado en *Anales CoNaIISI 2014*, eds: D. Riesco et al, Universidad Nacional de San Luis, ISSN: 2346-9927

² Facultad de Ingeniería e IESIING - Universidad Católica de Salta.

1. Introducción y motivación

En nuestra experiencia docente en la carrera de Ingeniería en Informática de la Universidad Católica de Salta hemos venido observando el interés de muchos alumnos por la robótica, algo atractivo pero prácticamente inaccesible en nuestro contexto. Es por eso que comenzamos a explorar la posibilidad de introducir el tema de manera extracurricular para los alumnos de ingeniería y para alumnos de colegios secundarios como elemento motivador. Dado el perfil de nuestra carrera, el énfasis estaría en la programación de robots, en lo posible de robots reales pero también, aprovechando la disponibilidad, de simuladores.

No es casualidad este interés por la robótica. Esta disciplina, que ciertamente no es nueva, se va haciendo cada vez más cercana por la extensión de las aplicaciones robóticas al ámbito del hogar y otras tareas de la vida diaria. Ya en 2007 Bill Gates predijo en un artículo en *Scientific American* titulado *Un robot en cada hogar* (gates, 2007), que la robótica sería la próxima área destacada de la informática. Por su interés, reproducimos aquí el primer párrafo de este artículo:

Imagínense que están presentes en el nacimiento de una nueva industria. Es una industria basada en tecnologías innovadoras, en la que un puñado de corporaciones bien establecidas vende dispositivos altamente especializados para usos industriales, y un número cada vez mayor de *startups* producen juguetes innovadores, gadgets para aficionados y otros productos interesantes. Pero también es una industria muy fragmentada, con pocos estándares o plataformas. Los proyectos son complejos, el progreso es lento, y las aplicaciones prácticas son relativamente raras. De hecho, a pesar de toda la excitación y la promesa, nadie puede decir con certeza cuándo – o si – esta industria llegará a tener masa crítica. Si lo consigue, seguramente podrá cambiar el mundo.

Al decir de Gates, este párrafo bien podría haberse referido a la industria de los PCs hace ya unas décadas, pero sin embargo está hablando de la industria robótica.

Según la *International Robotics Federation* (IRF, 2013), solo en 2012 se vendieron unos 3 millones de robots para uso personal y doméstico. La mayoría de ellos, 1,96 millones, se utilizan en tareas domésticas (limpieza del suelo, corte del pasto, etc.), con un incremento del 53% sobre 2011 y ventas de US\$697 millones. Los restantes son principalmente para entretenimiento (robots que son juguetes o para aficionados), educación e investigación. Las previsiones para el período 2013-2016 indicaban que se venderían unos 22 millones de unidades para uso personal, de los cuales 15 serían para tareas domésticas con un valor estimado de US\$5600 millones; unos 3,5 millones, para entretenimiento, unos 3 millones para educación e investigación, y unas 6.400 unidades para asistencia a personas ancianas y discapacitadas, un mercado, este último que aumentará sustancialmente en los próximos 20 años.

Así, cada vez está más cerca un futuro en que los robots y dispositivos relacionados formen parte de la vida diaria, gracias a tecnologías como la computación distribuida, el reconocimiento y síntesis de voz, el procesamiento de imágenes, y la conectividad inalámbrica de banda cada vez más ancha. Este futuro ciertamente motiva a muchos de nuestros alumnos, aun cuando no está exento de preocupaciones éticas.

Una empresa, iRobot, iniciada por investigadores del MIT, ha fabricado y vendido desde su

comienzo más de 10 millones de robots para uso doméstico, entre ellos su conocido modelo de robot aspiradora, Roomba (iRobot; 2014). Gracias a su éxito comercial, iRobot ha podido poner en el mercado un robot con la misma arquitectura que el Roomba para tareas de investigación y educación, el Create (Figura 1). Se trata de una variante del Roomba sin los accesorios de aspirar y cepillar. El precio del mismo de US\$150 (más US\$100 para una batería recargable), es un costo al alcance de las posibilidades de muchos interesados en la robótica, específicamente en la programación de robots.



Figura 1. Robot Create de iRobot con cámara inalámbrica utilizado en el taller.

Es importante destacar que la programación de los robots se está convirtiendo en el principal desafío, más que el hardware. Algunos investigadores de prestigio, como Red Whittaker y Sebastian Thrun, han afirmado que a pesar de las dificultades para que el hardware funcione adecuadamente, el avance en este campo es significativo y los escollos superables; ahora la dificultad está en el software que resuelva el problema (Gates, 2007; Thrun, 2006).

La existencia del Create, por su precio y calidad, nos motivó a poner en práctica el deseo de ayudar a nuestros alumnos a incursionar en el mundo de la robótica. Tras explorar el estado de la enseñanza de la robótica y las fortalezas y debilidades de nuestro equipo y contexto, y tras conseguir una unidad de este robot, avanzamos en la posibilidad de dictar talleres de introducción a la robótica, dejando la posibilidad abierta para conseguir en el futuro un número mayor de robots, en particular del tipo Lego Mindstorms, con mayor versatilidad desde su arquitectura física y variedad de sensores y actuadores. El Create viene listo para usarse, con una variedad de sensores, lo cual nos ha permitido centrarnos en la programación más que en el hardware. No obstante como plataforma el robot permite muchas posibilidades de extensión también desde el hardware y la electrónica.

Parte de nuestra exploración fue la búsqueda de alternativas para desarrollar software para robots, y específicamente para el Create y el Mindstorms. En base a ella decidimos inclinarnos por el Microsoft Robotics Development Studio (MRDS). Los aspectos en que basamos la decisión fueron la disponibilidad de un simulador de robots en diversos contextos; la posibilidad de controlar diversos robots usando la misma interfaz junto con la existencia de librerías para los dos robots mencionados, que nos permitieron comenzar a usar el Create de inmediato; la posibilidad de programar tanto con un lenguaje visual como con un lenguaje más tradicional, C#, conocido por buena parte de nuestros alumnos; y por último la disponibilidad del entorno en forma gratuita.

La introducción a la robótica, además de ofrecer a los estudiantes un acercamiento a la robótica, les permite aplicar conceptos y técnicas de programación que han ido adquiriendo durante la carrera, y ampliar sus conocimientos en otras que tal vez solo hayan visto desde la teoría. En particular, uno de los problemas más complejos de la programación en robótica es gestionar en simultáneo todos los datos provenientes de una diversidad de sensores, y emitir los comandos apropiados a los actuadores. En un enfoque de programación convencional un programa con un solo hilo (*thread*) iría sucesivamente leyendo datos de los sensores, después los procesaría y finalmente decidiría el comportamiento del robot, para volver a comenzar el ciclo de nuevo. Este enfoque tiene obvias limitaciones para la capacidad de respuesta ante imprevistos. Así, el alumno se enfrenta con el problema y gestión de la concurrencia, un desafío que se extiende más allá de la robótica a aplicaciones que precisen orquestar de manera eficiente un código que se ejecute, por ejemplo, en diferentes servidores o procesadores al mismo tiempo.

En resumen, hemos desarrollado el taller de introducción a la robótica, y más concretamente a la programación de robots, que describimos en este trabajo, con la intención de ayudar a los alumnos a acercarse al gran potencial, presente y futuro, de la robótica, combinando técnicas ya aprendidas por ellos a lo largo de su carrera y exponiéndolos a elementos teóricos y prácticos novedosos en esta interesante y excitante tecnología emergente.

En las dos próximas secciones describimos el entorno de desarrollo elegido y el enfoque y elementos del taller propuesto, para terminar este trabajo con un análisis de las ventajas e inconvenientes de la propuesta, líneas de trabajo futuro y algunas conclusiones.

2. MRDS

Microsoft Robotics Developer Studio (MRDS) es un entorno de programación basado en .NET para construir aplicaciones robóticas. En esta sección se describen los cuatro componentes principales de MRDS: CCR (*Concurrency and Coordination Runtime*), DSS (*Decentralized Software Services*), el lenguaje de programación visual VPL, y el entorno de simulación.

Como se ha mencionado, la concurrencia es un aspecto esencial en la programación robótica. Un enfoque para gestionar la concurrencia es escribir programas multihilo, siendo esta, una de las tareas más complejas en programación. MRDS esconde esta complejidad tras una librería llamada *Concurrency and Coordination Runtime* (CCR) para gestionar tareas asíncronas ejecutadas en paralelo, diseñada inicialmente para ayudar a los programadores a aprovechar las ventajas de sistemas multinúcleo y multiprocesador (Microsoft, 2010.a). Está escrita en C# y basada en .NET. CCR incluye una clase *Dispatcher* que implementa un pool de hilos, que

pueden ejecutarse simultáneamente. Cada despachador (*dispatcher*) consta de una cola de *delegados*, cada uno de los cuales representa el punto de entrada de una tarea que puede ejecutarse asincrónicamente. El despachador distribuye esas tareas entre los hilos. Cada elemento de la cola del despachador tiene un puerto, que es a su vez una cola en que se coloca la ejecución de la tarea. La tarea puede tener asociada un objeto *ReceiverTask* que consume el resultado para procesamiento posterior. Un árbitro gestiona este consumidor y lo invoca cuando el resultado esperado está disponible en el puerto.

Además de CCR, MRDS añade otro nivel denominado *Decentralized Software Services* (DSS) para combinar aplicaciones concurrentes en CCR que se denominan servicios, permitiendo además que los servicios corran en diferentes computadoras en una red. Se trata de un modelo de servicios *loosely coupled*. DSS organiza, crea y manipula servicios facilitando la creación de aplicaciones concurrentes y distribuidas. Los servicios – por ejemplo el código que lee un sensor o controla un motor – operan como procesos separados que pueden ser orquestados por otro servicio. El diseño de DSS sigue una arquitectura orientada a servicios web, en analogía con la forma en que texto, imágenes o información de diversos servidores se agrega en una sola página web. DSS utiliza el protocolo DSSP, que es similar a SOAP (Simple Object Access Protocol), utilizado para servicios web. La diferencia entre ambos es que mientras que el estado de un servicio web está oculto, el estado de un robot debe ser visible (el estado, por ejemplo, incluye la lectura de los sensores del robot, o la imagen capturada por su cámara). Así el estado de un servicio puede visualizarse u modificarse mediante un navegador web. DSSP separa el estado del servicio de su comportamiento u operaciones.

DSS permite que los servicios corran aisladamente, por lo que si un componente del robot falla, puede ser apagado o reiniciado desde la misma aplicación. Los servicios de un robot pueden también monitorearse o puede cambiarse su estado desde una ubicación remota usando la misma interfaz web.

Cuando se programa en MRDS la aplicación no reside en el robot, sino en una computadora (o distribuida entre varias). Así el robot puede ser un dispositivo relativamente sencillo que delega las tareas de procesamiento complejo (ej. visión, navegación y planificación de caminos con algoritmos sofisticados, orquestación de las tareas al nivel más alto) al hardware mucho más potente de la PC. Es posible incluso manejar grupos de robots que colaboren para realizar tareas complejas.

Las aplicaciones en C# pueden crearse utilizando la IDE Visual Studio, que tiene plantillas para crear un servicio que incluyen muchas de las directivas y elementos necesarios para el servicio. Hemos de decir que a pesar de estas ayudas la programación en este entorno es bastante compleja. Esta complejidad queda oculta en buena parte cuando se usa el lenguaje visual VPL, que se describe más adelante.

La vinculación entre el software y el hardware se define mediante archivos *manifest*; una misma aplicación puede correr en diferentes plataformas hardware cambiando simplemente el archivo *manifest*. Por ejemplo, un usuario podría programar usando comandos genéricos el movimiento de un dispositivo con tracción diferencial, luego y reemplazar el archivo *manifest* con el de cualquier robot, real o simulado, que utilice tracción diferencial. El servicio convierte automáticamente los comandos de alto nivel en instrucciones de bajo nivel específicas para ese tipo de robot. O también un servicio que precise una cámara, puede utilizar indistintamente

una cámara web, una cámara IP o la cámara del simulador simplemente cambiando el *manifest* y no el código del servicio. Esta es una ventaja de utilizar MRDS: la posibilidad de cambiar el hardware sin cambiar el código [6].

El tercer componente a destacar es un lenguaje de programación visual (VPL) y su correspondiente entorno de desarrollo. Está basado en un modelo de programación gráfica para el paradigma de flujo de datos, que modela un programa como un grafo dirigido de datos que fluyen entre operaciones. Un programa, más que una serie de comandos imperativos que se ejecutan en secuencia, es como un conjunto de trabajadores en una línea de ensamblado, que hacen su tarea tan pronto como reciben los componentes, en este caso los datos. Por ello VPL se adapta para escenarios de procesamiento concurrente y distribuido. Está dirigido a programadores principiantes, aunque otros más veteranos lo pueden aprovechar para el desarrollo rápido y prueba de prototipos de servicios. El código generado puede convertirse automáticamente en código en C# (Microsoft, 2010)

Por último, MRDS incluye un simulador, en que los usuarios pueden recrear entornos complejos y colocar en ellas robots. Hay una serie de escenas y de robots disponibles de antemano y un lenguaje para crear nuevas escenas y dispositivos simulados. El motor del simulador es capaz de simular colisiones, gravedad y otras fuerzas de manera muy aproximada a como los objetos se comportan en el mundo real. El uso del simulador facilita el acceso a la robótica a estudiantes aunque no se disponga de robots físicos suficientes.

La versión de MRDS que utilizamos es la R4 (2012) que precisa del sistema operativo Windows 7 u 8 y Visual Studio Express 2010 o superior. Con Windows XP solo puede utilizarsela versión R3 (2008). No existen versiones de MRDS para otros sistemas operativos. Todo el software está disponible de manera gratuita en la web o a través de la iniciativa académica de Microsoft, lo cual es una innegable ventaja en nuestro contexto. Como se indicó anteriormente, MRDS incluye librerías para utilizar el Create y el Lego Mindstorms, tanto con el robot real como en simulación.

3. Enfoque del taller

El taller diseñado consta de contenidos teóricos y prácticos, con marcado énfasis en este segundo aspecto. Surge como investigación de cátedra en el área de la inteligencia artificial (IA) y esto marca la orientación de la parte teórica hacia la IA y los robots inteligentes. El material introductorio de las clases teóricas está basado en el trabajo de Houston (2008) en cuanto a los elementos básicos de un robot, y en el de Murphy (2000) en cuanto a los diversos paradigmas de programación robótica (jerárquico, reactivo, e híbrido) y el proceso de diseño para construir robots inteligentes. Aunque no son obras muy recientes, son adecuadas para el nivel introductorio del taller. Para el trabajo de programación del Create con MRDS, hemos tomado como referencia la obra de Johns y Taylor (2008).

El núcleo del taller son los trabajos prácticos que los alumnos deben realizar, y que se irán describiendo en las secciones que siguen. En cada sesión se propone una actividad de programación facilitando a los alumnos un tutorial que indica los pasos a seguir; previo a ello se desarrolla un tema teórico vinculado con el práctico. En la elección de temas hemos adaptado parte de los propuestos por Warkman y Elzer (2008).

La mayor parte de los prácticos están diseñados para utilizar tanto el simulador como el robot real. MRDS permite esto con pequeños cambios en los programas creados por los alumnos, en particular con el intercambio de archivos *manifest*. Así se puede aprovechar mejor el recurso escaso que es el único robot Create disponible y los alumnos pueden continuar las prácticas en otro horario y lugar. En general se sugiere a los alumnos que trabajen en parejas, tanto por la disponibilidad del robot como por la ayuda que se pueden prestar en el aprendizaje, ya que algunas de las tareas son difíciles de implementar y de hacer funcionar en VPL, especialmente para principiantes; esto hace que el trabajo sea en ocasiones muy satisfactorio y en otras bastante frustrante.

Aunque nuestros alumnos tienen experiencia en programar en diversos lenguajes de .NET, incluyendo C#, la curva de aprendizaje de CCR y DSS es demasiado pronunciada para un taller de las características que nos proponemos. Por ello en el taller se utiliza el lenguaje VPL, aunque al final se propone a los alumnos el desafío de programar servicios en C# y se les proporcionan algunos materiales a modo de tutorial.

El hardware disponible para los prácticos incluye además del robot Create con su batería recargable, una cámara web inalámbrica, y un módulo BAM para conectar el robot mediante Bluetooth a las computadoras en que trabajan los alumnos (estas deben disponer de conexión Bluetooth o utilizar el *dongle* USB y software del que dispone el taller). También se dispone de una pared virtual diseñada específicamente para el robot Create.

3.1. Práctico I: Introducción a la robótica, al MRDS y al robot Create

El primer práctico pide a los alumnos que realicen los tres primeros tutoriales proporcionados por los creadores del MRDS y disponibles en la web (traducidos por nosotros). El primer tutorial introduce el concepto de servicios usando VPL. Los servicios son el núcleo básico de la programación robótica en nuestro entorno. Un servicio es una entidad (implementada con orientación a objetos en C#) que controla un componente hardware o software particular. Como se indicó, una aplicación robótica consiste en orquestar una serie de servicios, y combinando servicios muy sencillos se puede obtener un comportamiento complejo. En el segundo tutorial se combinan varios servicios, en particular un sensor de impacto y la tracción diferencial del robot. Los alumnos aprenden a encender y apagar los motores del robot mediante un botón, y pueden extender ese comportamiento para controlar la velocidad del robot, por ejemplo.

El tercer tutorial añade algo más de funcionalidad con los mismos elementos, haciendo que el robot deambule y, si choca con un obstáculo retroceda una pequeña distancia, gire aleatoriamente por un corto periodo de tiempo, y luego avance en la nueva dirección. Además se propone construir una nueva actividad en VPL (análoga a una subrutina) que nuclea este comportamiento. Como ya se indicó, el mismo código puede ser utilizado tanto en el simulador (para lo cual proporcionamos un escenario con un robot y diversos obstáculos, mostrado en la Figura 2) como con el robot real, simplemente modificando el archivo *manifest* utilizado. El objetivo de este práctico es familiarizarse con el funcionamiento de MRDS y VPL, en particular crear un proyecto, acceder a los servicios, y conectarlos entre sí en el paradigma de flujo de datos que sigue VPL. En la introducción a los tutoriales se explica brevemente la diferencia entre la programación basada en flujo de datos y la basada en flujo de control, que es a la cual los alumnos

están acostumbrados. Este es un aspecto que suele causar dificultades conceptuales en cuanto la complejidad de los programas va aumentando.

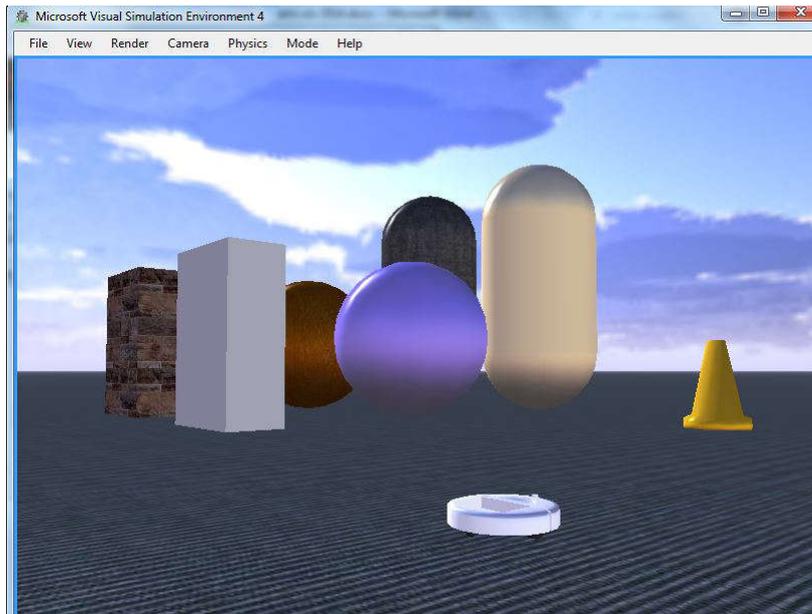


Figura 2. Escenario de simulación por defecto para el robot Create.

El práctico en sí no tiene mayor dificultad ya que los alumnos van siguiendo los pasos especificados en los tutoriales. No obstante, en base a nuestra experiencia con estas herramientas, este trabajo puede llevar tiempo considerable y encontrar dificultades complejas y sutiles para hacer que el código funcione.

3.2. Práctico II: Teleoperación

El segundo práctico toma como punto de partida un tutorial sobre teleoperación también disponible en MRDS, en el que además se debe aplicar lo aprendido en el primer práctico. El material teórico cubre el tema de teleoperación. Aunque nuestro énfasis es en robots autónomos, este es un tema importante, un punto intermedio necesario entre la mayoría de los robots industriales, programados para realizar tareas de precisión en entornos predecibles y no cambiantes, y los robots inteligentes que deben desenvolverse en entornos desconocidos y cambiantes y responder adecuadamente, sin poder ser programados de antemano para responder y adaptarse a todas las posibles contingencias (Murphy, 2000). El puente entre ambos enfoques tradicionalmente han sido mecanismos que permiten que una persona controle de manera remota el robot, o partes del mismo. A este conjunto de mecanismos, locales o remotos, suele llamarse teleoperación. Además la teleoperación forma parte de muchas de las nuevas aplicaciones de robots para la telepresencia, un conjunto de tecnologías que permiten a una persona

sentir como si estuviera presente en un lugar remoto, o dar la impresión a otros de que lo está (Lehrbaum, 2013).

En este práctico los alumnos crean una interfaz para controlar el robot remotamente y visualizan la imagen devuelta por la cámara ubicada en el robot. Así navegan un recorrido con obstáculos, pero no viendo el robot sino la imagen obtenida por la cámara montada sobre el mismo, como suele ocurrir en la mayoría de las aplicaciones de teleoperación. El práctico incluye una competencia para ver qué equipo de alumnos es más hábil en teledirigir el robot y evitar los obstáculos, algo más difícil de lo que parece.

El primer práctico constaba de sencillos ejemplos. En este se construye en base a ellos una aplicación práctica. El carácter competitivo añade motivación para que los alumnos vayan un poco más allá de lo propuesto en el tutorial, especialmente en cuanto a la interfaz de teleoperación.

3.3. Práctico III: Deambular y cubrir un recinto

Una de las tareas frecuentes de un robot es deambular recorriendo su entorno. En la instancia teórica de esta actividad, se presenta a los estudiantes la importancia de este comportamiento y ciertos enfoques al mismo (Murphy, 2000) y luego, en la práctica, los mismos pueden implementar sus propios algoritmos, en particular para conseguir que el Create cubra el suelo de la habitación de la manera más eficiente posible. Después deben añadir comportamiento que haga que el robot evite obstáculos usando los sensores disponibles en el Create (sensor de impacto y sensor de infrarrojo que detecta una pared cercana). Así pueden comprobar la efectividad, ventajas e inconvenientes de diversos tipos de sensores para una tarea específica. No existe un sensor perfecto, sino que deben aprovecharse los sensores disponibles en el robot. Tampoco puede decirse que los enfoques más complejos siempre realicen la tarea mejor que los comportamientos muy sencillos.

Hay una relación entre este práctico y una aplicación real, ya que la tarea propuesta es la básica del robot aspirador Roomba, mellizo de nuestro robot Create, con los mismos sensores y actuadores. El éxito del Roomba está en gran medida basado en la efectividad de sus algoritmos para deambular y recorrer un recinto.

3.4. Práctico IV: Ejemplo del paradigma reactivo

La tarea de este práctico está inspirada en los comportamientos biológicos para buscar comida y evitar peligros, tales como depredadores. Con esta idea se pretende mostrar los fundamentos del paradigma reactivo en robótica. Los alumnos tienen libertad para elegir cómo implementar esos comportamientos. Para ello deben utilizar la cámara web inalámbrica disponible o el simulador junto con un servicio que realiza un procesamiento de imágenes sencillo para detectar «comida» -pelotas de color azul- y «depredadores» -pelotas de color rojo (ver Figura 3). Este servicio ha sido desarrollado por nosotros usando AForge.NET, un framework en C# para el desarrollo de aplicaciones de procesamiento de imágenes, redes neuronales, algoritmos genéticos, robótica, etc (www.aforgenet.com).

Dado que nuestro robot no tiene un brazo, lo que los alumnos pueden realizar con comida y depredadores es bastante limitado. No obstante hay una variedad de respuestas que pueden

implementar, por ejemplo alejarse inmediatamente del peligro o tratar de rodearlo como si fuera un obstáculo.



Figura 3. Entorno para el práctico IV.

3.5. Práctico V: Seguidor de línea

En este práctico los alumnos enfrentan un ejemplo típico de robótica: seguir una línea marcada en el suelo. En nuestro caso se utilizan los sensores de abismo del Create, cuyo uso original es detectar desniveles, tales como escaleras. Estos sensores utilizan infrarrojos, por lo que pueden aprovechar las diferencias en las propiedades de reflexión de distintas superficies para detectar la ausencia o presencia de la línea. Así los alumnos aprecian la versatilidad, ventajas e inconvenientes de diversos sensores. Este práctico permite el trabajo tanto con el robot real como con el simulado; en el segundo caso se aprovechan los servicios y los escenarios del simulador proporcionados por Lehrbaum (2013).

3.6. Práctico VI: Contar objetos

Este último práctico, cuyo objetivo es contar los objetos existentes en un entorno, es opcional; depende de cuánto cada alumno o equipo avance con los prácticos anteriores, ya que los conocimientos adquiridos facilitan esta tarea, que por otro lado no está exenta de dificultades, como la de saber qué objetos ya han sido contados.

Una alternativa es dejar libertad a los alumnos para atacar un problema que sea de su interés después de todas las experiencias anteriores. Aunque el hardware disponible, especialmente la variedad de sensores, es limitado, hay muchas posibilidades de temas que no son abordados en el tiempo disponible. Así los integrantes del taller, pueden integrar los conceptos anteriores y explorar algo de su propio interés.

4. Ventajas e inconvenientes del diseño del taller

En el diseño del taller se intentó hacer corresponder temas básicos teóricos sobre robótica con los trabajos prácticos. Así los alumnos pueden comprender algunos de los problemas principales de la robótica, especialmente desde el punto de vista de la programación, y aplicar en el laboratorio alguno de los principios básicos para resolverlos.

Por tratarse de un taller opcional para alumnos interesados, y no un curso formal del currículum, no se puso el énfasis en diseñar actividades de evaluación. Es difícil cuantificar la creatividad y el ingenio a la hora de resolver alguno de los prácticos. Además no siempre los esfuerzos se traducen en el comportamiento adecuado o hasta deseado del robot, lo cual puede ser resultado de las limitaciones del hardware (ej. precisión de los sensores o actuadores) y del software disponible (ej. para procesado de imágenes, o las características del lenguaje y entorno VPL).

Desde el comienzo MRDS pareció una alternativa adecuada para un taller de estas características. En cierto modo proporciona un puente entre lo que los alumnos ya conocen (sintaxis de lenguajes de programación, o hasta el entorno Visual Studio y .NET) y los elementos semánticos nuevos propios de la robótica. Otra ventaja clara fue la posibilidad de comenzar a trabajar inmediatamente con el robot disponible y con un simulador gracias al soporte y librerías incluidas en la herramienta. Esta es una ventaja importante dado que nuestro énfasis, como ya se indicó, es en la programación robótica y no tanto en la electrónica u otros aspectos de hardware de bajo nivel. La calidad y realismo del simulador, la variedad de escenarios proporcionados y la posibilidad de diseñar nuevos escenarios con relativa facilidad son aspectos muy atractivos. La disponibilidad de tutoriales y otros materiales para iniciarse en el entorno MRDS es otra ventaja adicional. La disponibilidad gratuita de las herramientas necesarias fue también un factor determinante.

Sin embargo la curva de aprendizaje para nuestro equipo ha sido bastante pronunciada. Una de las tareas arduas de la preparación del taller ha sido familiarizarnos lo suficiente con el software para poder transmitir conceptos y tecnología de manera adecuada. En el proceso hemos encontrado muchas dificultades, tales como que algunos aspectos del entorno no funcionaban como se esperaba, errores del software, falta de material en castellano y la necesidad de aprender diversos tópicos necesarios, desde la programación orientada a servicios con CCR y DSS, hasta las dificultades propias de los servicios concretos para controlar el robot.

Microsoft lanzó la iniciativa de MRDS como una propuesta de una plataforma estándar para programar robots; la versión más reciente es de marzo de 2012 e incluye amplio soporte para el dispositivo Kinect. Sin embargo no está claro el futuro de esta herramienta, el sitio web no ha variado sustancialmente desde entonces, y el interés y soporte en los foros disponibles parece ir disminuyendo (aunque las preguntas que realizamos en los mismos fueron contestadas relati-

vamente pronto por los desarrolladores de la herramienta). Uno de los factores que podrían causar el desinterés de su uso, es su complejidad, que es notable cuando se trata de extender la herramienta a otros sensores o robots diferentes de los muchos disponibles inicialmente. En teoría parece una tarea sencilla dada la cantidad de ejemplos y de implementaciones que se pueden adaptar a nuevos robots; en la práctica la complejidad es grande.

No obstante, el escenario de trabajo hardware y software que hemos descrito nos ha dado ya la posibilidad de introducir la robótica de manera accesible en nuestro contexto y con capacidad de entusiasmar a nuestros alumnos y aumentar su interés y curiosidad.

5. Trabajo futuro

A la hora de redactar este trabajo estamos a punto de ofrecer este taller por primera vez a un grupo de alumnos. La evaluación de esta experiencia, que esperamos repetir regularmente, nos dará información para trazar el camino a seguir en el futuro. Idealmente podremos aumentar el número y variedad de robots disponibles. Como se mencionó, una de las posibilidades es utilizar el Lego Mindstorms. El taller no precisaría grandes modificaciones, pero los sensores disponibles (ej. sónar) y las posibilidades de configuración del robot (por ej. como un brazo robótico) amplían el rango de posibilidades de lo que los asistentes al taller podrían llegar a diseñar y construir a lo largo del mismo.

Varios de los alumnos están comenzado proyectos de grado con la tecnología presentada y aumentándola en diversas direcciones. Uno de los aspectos a explorar es el uso de otras herramientas software que permitan el desarrollo de proyectos más avanzados, incluso en el ámbito de la investigación, y que puedan reemplazar el uso de MRDS, especialmente si éste cae en desuso o se torna obsoleto.

6. Conclusiones

La disponibilidad de un robot y especialmente de un equipo de personas interesadas en la robótica ha generado muchas expectativas ente los alumnos y profesores de la Facultad. Aunque el taller propuesto puede mejorarse en diversos aspectos, cumple las expectativas de motivar a nuestros alumnos a incursionar en la robótica y de exponerlos a esta tecnología tan innovadora en nuestro contexto pero de tanta importancia para el futuro de sus carreras y hasta de su estilo de vida.

La disponibilidad de robots de bajo costo, como el Create de iRobot, y de software de acceso gratuito para la programación de robots han hecho posible el acceso de nuestros alumnos a esta tecnología. El campo está abierto además para que los alumnos puedan llevar a cabo trabajos finales e investigaciones a nivel de cátedra, lo que genera entre nosotros grandes expectativas.

Referencias

Gates, B. «A robot in every home» *Scientific American*, Enero 2007: 58-65.

Housten, D. «Learning Roomba». MSc Thesis, Drexel University, 2008.

iRobot, «Our History,» 2014. Disponible en http://www.irobot.com/us/Company/About/Our_History.aspx. [Último acceso: 11 07 2014].

Johns, K. y T. Taylor. «Professional Microsoft Robotics Developer Studio». Wiley, 2008
Lehrbaum, R. «Attack of the Telepresence Robots!,» *Information Week*, pp. Disponible en <http://www.informationweek.com/applications/attack-of-the-telepresence-robots!/d/d-id/1108137>, 1 11 2013.

Microsoft. «Microsoft Robotics User Guide» 2010. [En línea]. Disponible en <http://msdn.microsoft.com/en-us/library/dd936006.aspx>.

Microsoft. «Visual Programming Language,» 2010a. Disponible en <http://msdn.microsoft.com/en-us/library/bb964572.aspx>.

Murphy, R. «Introduction to AI robotics». Cambridge, Massachusetts: The MIT Press, 2000.

RF.«World Robotics - Service Robots 2013. Executive Summary,» Disponible en http://www.worldrobotics.org/uploads/tx_zeifr/Executive_Summary_WR_2013_01.pdf, 2013.

Thrun, S. «The great robot race,» *NOVA - PBS*, p. Disponible en http://www.pbs.org/wgbh/nova/transcripts/3308_darpa.html, 28 Marzo 2006.

Workman, K. and S. Elzer. «Creating An Upper-Level Undergraduate Robotics Elective in Computer Science» de *Proceedings of the 23rd Annual Conference of the Pennsylvania Computer and Information Science Educators (PACISE)*, Kutztown, PA, 2008.

Seguridad de las Redes de Control Industrial – MODBUS/TCP con inspección profunda de paquetes¹

Carlos Said²

Resumen

El artículo propuesto tiene como objetivo tratar la aplicación y la gestión de las prácticas de seguridad cibernética asociadas con la tecnología de la información y la tecnología operativa en relación con los entornos de redes de producción industrial (SCADA-ICS). Se ahonda en los aspectos de las tecnologías de inspección profunda de paquetes de información y granularidad en el análisis de protocolos de datos, también llamados DPI.

Palabras clave: protocolos – SCADA - ICS, granularidad - DPI

Abstract

The proposed article aims to address the implementation and management of cyber security practices associated with information technology and operational technology regarding network environments of industrial production (SCADA-ICS). The guidance in this article is to complement comprehensive cybersecurity specific areas related to deep packet inspection or granularity analysis of transmission protocols.

Keywords: protocols, SCADA, ICS, granularity, DPI

Introducción

Puede observarse que el aumento del riesgo de la seguridad cibernética en redes SCADA (Supervisory Control and Data Acquisition, en español Redes de Supervisión, Control y Adquisición de Datos) / ICS (Industrial Control Systems, en español Sistemas de Control Industrial) lleva a los proveedores de tecnologías de seguridad, a desarrollar mecanismos avanzados que se ocupen de los desafíos específicos de seguridad de los sistemas de control.

Una manifestación tangible de ello, se observa en el gráfico 1. Este identifica vulnerabilidades de seguridad explotadas en redes de control industrial.

¹ Este artículo fue presentado en: 2nd PAN-American Interdisciplinary Conference, PIC 2016, 24-26 February 2016, Buenos Aires, Argentina

² Facultad de Ingeniería - Universidad Católica de Salta.

ID	Disc Date	Title
117992	2015-01-30	SCADA SpecView Unspecified Remote Information Disclosure
		SCADA SpecView contains an unspecified flaw that may allow a remote attacker to gain access to potentially sensitive information. No further details have been provided.
119930	2015-03-13	Events SCADA HMI Unspecified Information Disclosure
		Events SCADA HMI contains an unspecified flaw that may allow an attacker to gain access to potentially sensitive information. No further details have been provided by the researcher.
119492	2015-03-10	SCADA Engine BACnet OPC Server BACnetOPCServer.exe Format-String
		SCADA Engine BACnet OPC Server contains a format string flaw in BACnetOPCServer.exe. The issue is triggered as format string specifiers (e.g. %s and %x) are not properly sanitized in user-supplied input when handling a specially crafted request. This may allow a remote attacker to cause a denial of service or potentially execute arbitrary code.
119406	2015-03-10	SCADA Engine BACnet OPC Server Remote Item Manipulation
		SCADA Engine BACnet OPC Server contains a flaw related to authentication that may allow a remote attacker to insert, read, or delete any items in the database. No further details were provided.
120994	2015-04-17	Modbus SCADA (WLC Systems) Unspecified Traversal Remote Issue
		Modbus SCADA (WLC Systems) contains a flaw that allows traversing outside of a restricted path. The issue is due to the program not properly sanitizing user input, specifically path traversal style attacks (e.g. '..'). With a specially crafted request, a remote attacker can have an unspecified impact.
117990	2015-01-30	ScadaBR Unspecified File Upload Remote Command Execution
		ScadaBR contains an unspecified flaw that allows a remote attacker to execute arbitrary commands. This flaw exists because the program does not properly verify or sanitize user-uploaded files. By uploading a crafted file, the remote system will place the file in a user-accessible path. Making a direct request to the uploaded file will allow the attacker to execute commands with the privileges of the web server.
117102	2015-01-13	Clorius Controls ISC SCADA Java Web Client Insecure Credential Encryption MITM Credential Disclosure
		Clorius Controls ISC SCADA contains a flaw in the Java web client. The issue is triggered when credentials are not sufficiently encrypted, which may disclose credentials to a man-in-the-middle attacker.
119431	2015-03-10	SCADA Engine BACnet OPC Server Packet Handling Remote Heap Buffer Overflow
		SCADA Engine BACnet OPC Server contains an overflow condition that is triggered as user-supplied input is not properly validated when handling a specially crafted packet. This may allow a remote attacker to cause a heap-based buffer overflow, resulting in a denial of service or potentially allowing the execution of arbitrary code.

Gráfico 1. Incidentes de seguridad en redes SCADA-ICS

Fuente: <http://blog.osvdb.org/category/vulnerability-statistics/>

Es prudente iniciar el trabajo haciendo mención a que SCADA hace referencia a un sistema que opera sobre canales de comunicación para brindar control de equipos remotos.

El sistema de control puede combinarse con un sistema de captura/adquisición de datos para que a través del canal de comunicación se conozca el estado del equipo remoto, con fines de monitoreo, registro y/o control. En este caso suele llamarse sistema de control industrial (ICS). Los procesos industriales que se desean controlar son procesos que existen en el mundo físico. Los sistemas SCADA suelen distinguirse respecto de otros sistemas ICS por el hecho de controlar procesos de gran escala, que incluyen dispersión geográfica (número de sitios) y distancias relevantes.

Un aspecto de estos desafíos es el uso generalizado de protocolos de comunicación para sistemas ICS que no fueron diseñados pensando en la seguridad. Asegurar estos protocolos sin afectar su funcionalidad de control requiere de tecnología avanzada.

¿Cuáles son estos protocolos? MODBUS/TCP, BACnet IP, DNP3, solo para citar algunos de ellos.

Sin ahondar en este artículo, en todos estos protocolos (nuestro foco estará en MODBUS), pero debemos al menos mencionar en qué consisten los dos restantes que hemos mencionado).

BACnet es un protocolo de comunicación de datos para la construcción de redes de automatización y control. Fue desarrollado por motivación de de la Sociedad Americana de Ingenieros de Calefacción, Refrigeración y Aire Acondicionado (ASHRAE).

BACnet es un estándar nacional norteamericano, una norma europea, una norma nacional en más de 30 países, y un estándar mundial ISO. El protocolo es mantenido por el Comité Permanente nº. 135 del Proyecto para el Standard ASHRAE Standard.

DNP3 (Distributed Network Protocol, en español Protocolo de Red Distribuido) es un conjunto de protocolos de comunicación utilizados entre los componentes de los sistemas de automatización de procesos. Su uso principal es en empresas de electricidad y agua. Fue desarrollado para las comunicaciones entre los distintos tipos de equipos de adquisición y control de datos. En los sistemas SCADA, se utiliza en las estaciones maestras (master o centros de control), unidades terminales remotas (en inglés RTU, Remote Terminal Units), y dispositivos electrónicos inteligentes (en inglés, IED, Intelligent Electronic Devices). Se utiliza esencialmente para las comunicaciones entre una estación maestra y las RTUs o IEDs.

ICCP, (en inglés inter-control center communications protocol), es una parte de la norma IEC 60870-6, el cual se utiliza para las comunicaciones entre las estaciones maestras.

Retornemos a los aspectos de seguridad. Un ejemplo de estos mecanismos avanzados de seguridad de la información que ‘fluye’ usando estos protocolos como transporte es la inspección profunda de paquetes (en inglés, DPI o Deep Packet Inspection).

Por un lado, los sistemas de detección de intrusiones (IDS, Intrusion Detection Systems) brindan prestaciones de monitoreo para categorías genéricas de ataques básicos. Por otro lado, los firewalls utilizan básicamente listas de control de acceso, que permiten o bloquean todos los mensajes de un protocolo industrial como Modbus TCP.

Protocolo MODBUS

MODBUS es un protocolo de comunicación serial publicado originalmente por Modicon (ahora Schneider Electric) para su uso con sus controladores lógicos programables (en inglés PLC, programmable logical controllers). Es simple y por ello se convirtió desde entonces en un protocolo de comunicación estándar de facto, disponible para la conexión de dispositivos electrónicos industriales. Fue desarrollado con aplicaciones industriales en mente, es público y gratuito, y esencialmente ‘transporta bits’ sin imponer mayores restricciones.

Modbus permite la comunicación entre múltiples dispositivos conectados a la misma red: por ejemplo un sistema que mide la temperatura y la humedad y comunica los resultados a un computador. Modbus se utiliza a menudo para conectar un computador supervisor con una terminal remota (RTU) en redes SCADA. Los tipos de datos suelen nombrarse con base en el uso en los relés de conducción: una salida física de un solo bit se llama bobina (o en inglés, Coil), y una entrada física de un solo bit se llama ‘entrada discreta’ o ‘contacto’.

La actualización de los protocolos Modbus está a cargo de la Organización Modbus. Es una asociación de usuarios y proveedores de dispositivos compatibles con Modbus que busca impulsar la evolución del protocolo.

Solemos distinguir entre Modbus RTU sobre TCP (en inglés transmission control protocol) y de Modbus TCP.

Modbus RTU sobre TCP es en síntesis, un mensaje Modbus RTU transmitido con una envoltura (o empaquetado) TCP/IP y se envía a través de una red en lugar de líneas seriales. El servidor no tiene ‘SlaveID’ (identificación segura) ya que utiliza en lugar de ello una dirección IP.

Respecto a Modbus TCP, la guía de implementación de mensajes Modbus proporcionada por Schneider Automation define un protocolo modificado específicamente para su uso a través de TCP/IP. La especificación oficial Modbus se puede encontrar en www.modbus.org/specs.php. Las principales diferencias entre Modbus RTU y Modbus TCP escapan al alcance de este artículo, siendo eventualmente el objeto de una próxima redacción.

La estructura de un mensaje Modbus IP está dividido en dos partes. La primera de ellas es llamada Encabezamiento del Protocolo de Aplicación Modbus (o en Inglés, Modbus Application Protocol (MBAP) Header), que consiste de:

1. Identificación de la transacción (2 bytes) – Usado para sincronización entre mensajes cliente-servidor.
2. Identificación del protocolo (2 bytes) – Se encuentra en 0 para Modbus, y se reserva para usos futuros.
3. Longitud (2 bytes) – Usado para definir la cantidad de bytes restantes en la estructura del mensaje.
4. Identificador de la unidad (1 byte) – Dirección de la ‘unidad esclava’ a la cual le está siendo enviado el mensaje. Para Modbus, la dirección de del dispositivo / unidad ‘esclavo’ es la dirección IP, y por ello el identificador de unidad es puesto al valor hexadecimal 0xFF.

La segunda parte del mensaje es típicamente un mensaje de datos del protocolo Modbus (en Inglés, Modbus Protocol Data Unit (PDU)). Consiste de:

1. Código de Función (1 Byte) – Funciones soportadas por Modbus.

2. Datos de la función (n bytes) – Datos que acompañan el código de función, tales como respuestas o comandos.

Es esencialmente en esta segunda parte del mensaje donde las herramientas y técnicas de inspección profunda de paquetes, realiza su tarea de análisis.

Encabezamiento del Protocolo de Aplicación Modbus				Mensaje de datos del protocolo Modbus	
Identificación de la transacción	Identificación del protocolo	Longitud	Identificador de la unidad	Código de Función	Datos de la función

Gráfico 2. Estructura de un mensaje Modbus IP

La granularidad en el análisis de los mensajes

La inspección profunda de paquetes, ahonda en el análisis de los datos, para intentar entender para qué se utiliza el protocolo, y como consecuencia de ello proporcionar protección, no sólo detección. Para hacer esto debe tener la capacidad de ‘mirar’ y analizar los campos del mensaje de datos del protocolo Modbus (ver gráfico 2).

Para realizar lo mencionado en el párrafo previo, DPI puede por ejemplo, mediante la determinación de si un mensaje Modbus es de lectura o de escritura, descartar todos los mensajes de escritura; o sólo permitir la escritura de determinados registros. Esto da lugar a que la protección se adapte a las necesidades de la aplicación, permitiendo que los mensajes de control esenciales fluyan según sea necesario, mientras que se bloquean los mensajes potencialmente peligrosos o inapropiados.

Para entender DPI, primero es importante entender cómo funciona un firewall tradicional. Un firewall es un dispositivo que monitorea y controla el tráfico que fluye en una red o entre redes. Captura el tráfico que pasa a través de él y lo compara con un conjunto predefinido de reglas (llamado Access Control Lists o ACL, Lista de Control de Accesos). Cualquier mensaje que no coincide con las ACL se descarta.

El firewall clásico permite que las ACL puedan comprobar tres campos principales en un mensaje, a saber:

1. La dirección IP de la computadora que envía el mensaje (IP de origen),
2. La dirección IP del equipo que recibe el mensaje (IP de destino),
3. El protocolo de la capa superior que figura en el paquete IP, como se define en el campo «Número de puerto TCP de destino».

El tercer punto, número de puerto de destino TCP puede necesitar un poco más de explicación. Estos puertos no son puertos físicos, como un puerto de Ethernet; son números especiales que forman parte de cada mensaje TCP o UDP para identificar el protocolo de la capa de aplicación que se transporta en el mensaje.

Por ejemplo, Modbus/TCP utiliza el puerto 502. Estos números están registrados bajo la autoridad regulatoria Internet Assigned Numbers Authority (IANA).

Para poner todo lo dicho en conjunto, imagine que sólo desea permitir que el tráfico de Internet (el tráfico HTTP, en inglés Hyper Text Transfer Protocol) de un cliente en la dirección IP 192.168.1.20, hacia un servidor web con una dirección 192.168.1.51. Entonces escribiría una regla ACL similar a la siguiente

Permitir Origen=192.168.1.20 Destino=192.168.1.51 Puerto=HTTP

Se podría cargar esta ACL en el firewall y siempre y cuando se cumplan los tres criterios, el mensaje será permitido.

O supongamos que desea bloquear todo el tráfico Modbus que circule a través del firewall. Se podría definir una regla que bloquea todos los paquetes que contienen 502 en el campo de puerto de destino.

El problema con este esquema es su dicotomía. Permite un determinado protocolo o lo bloquea. No es posible el control de ‘grano fino’, la granularidad en el análisis, del protocolo.

Esto no es lo mejor, porque los mismos protocolos SCADA/ICS no tienen granularidad. Desde el punto de vista del número de puerto, un mensaje de lectura de datos se ve similar a un mensaje de actualización del firmware (o de escritura).

Por esto si permitimos que mensajes de lectura de datos, desde un HMI (en inglés, Human Machine Interface, o Interfaz Hombre-Máquina) a un PLC, pasen a través de un firewall tradicional también se está permitiendo que los mensajes de programación (o escritura) pasen a través de él. Se trata de un problema de seguridad relevante.

El firewall tiene que ahondar en los protocolos para entender exactamente para que está siendo utilizado. Una vez se aplican las reglas tradicionales, el firewall inspecciona el contenido de los mensajes y aplica reglas más detalladas.

Por ejemplo, un firewall Modbus con DPI determina si el mensaje Modbus es un mensaje de lectura o un mensaje de escritura y luego descarta todo mensaje de escritura.

Los firewalls con DPI pueden también «sanitizar» (expresión usada para manifestar: ‘transformar en seguro’) el tráfico de mensajes con formato o comportamientos inusuales.

Ejemplo

Un caso real de una empresa responsable de la gestión de puertos. Un control detallado/granular (también llamado ‘control de grano fino’) del tráfico SCADA/ICS mejora la seguridad y confiabilidad de uno de sus sistemas. La misma utiliza PLCs en sus controles y puentes para garantizar la seguridad de las naves marítimas y el tráfico de vehículos. Asegurarse de que estos PLC no sean manipulados es crítico para la seguridad tanto de las embarcaciones como del tráfico vehicular en los puentes.

El problema que esta empresa enfrenta es que un número de computadoras necesita ejecutar continuas operaciones de acceso a los datos de los PLCs. Sin embargo sólo a los equipos de control especiales se debe permitir enviar comandos y afectar el funcionamiento de los mismos. El mecanismo de contraseña tradicional o soluciones de firewall estándares no se consideran seguros, porque no ofrecen el control de ‘grano fino’ o granularidad en el análisis del mensaje, necesario.

La solución fue utilizar firewalls Modbus DPI para controlar todo el tráfico a los PLCs. Sólo a los mensajes Modbus de lectura se les permite llegar a los PLCs (a excepción de algunos equipos de alta seguridad). Todos los comandos de programación Modbus remotos (es decir de escritura, no de lectura) son bloqueados por lo que la programación se limita a los ingenieros ubicados en los sitios.

Proporcionar seguridad tan solo bloqueando o permitiendo familias completas de protocolos entre las redes no es suficiente para las operaciones SCADA/ICS actuales. Los protocolos en los cuales estos sistemas se sustentan son de gran capacidad funcional y potencialmente expuestos a ataques cibernéticos. Es tiempo de considerar cómo podemos utilizar las tecnologías DPI para que nuestros sistemas de control industrial sean más seguros y confiables.

Conclusiones

Los analistas de la industria mencionan la existencia de más de siete millones de nodos MODBUS, tan solo en Estados Unidos y Europa. Por esto se torna relevante el análisis de los esquemas de seguridad implementados. Sirva tan solo a modo de referencia la evaluación de los siguientes aspectos al considerar un firewall como esquema de seguridad. No es la única evaluación factible y cada industria/empresa deberá realizar su propia evaluación, inclusión y exclusión de factores a la hora de escoger.

Funcionalidad	Packet Filter Firewalls	Stateful Firewalls	Aplicación	Deep Packet Inspection Firewall industrial
1. Definir reglas básicas sin considerar relaciones entre paquetes	SI	SI	SI	SI
2. Definir reglas básicas considerando relaciones entre paquetes	NO	SI	SI	SI
3. Definir reglas básicas a nivel aplicación (DPI).	NO	NO	Dep	SI
4. Bloquear tráfico con malware al poder segmentar por <u>protocolos industriales</u>	NO	Dep.	NO	SI
5. Definir reglas a nivel protocolo (function codes) si el protocolo es <u>Modbus, Ethernet/IP</u> o asignar un único puerto sobre <u>OPC</u>	NO	NO	NO	SI
6. Eventos y logs	SI	SI	SI	SI
7. Preparados para entornos industriales	NO	NO	NO	SI

Fuente: <http://www.ciberseguridadlogitek.com/>

Referencias

Good Practice Guide on Firewall Deployment for SCADA and Process Control Networks, Centre for the Protection of National Infrastructure (CPNI), London.

Morris, T. H., Jones, B. A., Vaughn, R. B. & Dandass, Y. S. (2013), Deterministic Intrusion Detection Rules for MODBUS Protocols, in «Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS), 2013», IEEE, pp. 1773-1781.

NIST SP: 800-12. An Introduction to Computer Security: The NIST Handbook.

NIST SP 800-82 Rev 2, Guide to Industrial Control Systems (ICS) Security, May 2015.

Programmable Logic Controllers. A Practical Approach TO IEC 61131-3 Using CoDeSys, Autor: Dag H. Hanssen, Institute of Engineering and Safety, University of Tromso, Norway. Editorial: John Wiley & Sons, Ltd, 2015.

www.bacnet.org

www.modbus.org/specs.php

www.simplymodbus.ca/TCP.htm

Perfil académico de los autores

ALBARRACÍN, CARLOS MARCELO

Ingeniero Mecánico Aeronáutico egresado de la Universidad Nacional de Córdoba. Doctor en Ingeniería de la Universidad Nacional de Salta. Se desempeña como Profesor Adjunto, responsable de la cátedra de Análisis Numérico de la carrera de Ingeniería Química de la Facultad de Ingeniería de la Universidad Nacional de Salta.
cmalba@unsa.edu.ar

ALBARRACÍN, FERNANDO JAVIER

Ingeniero Civil egresado de la Universidad Católica de Salta y Miembro del Instituto Interdisciplinario de Ingeniería (I. Es. I. ING) de la Facultad de Ingeniería de la Ucasal. Se desempeña como Profesor Adjunto de la cátedra de de Estabilidad III de la carrera de Ingeniería Civil de la Facultad de Ingeniería de la Ucasal, y como jefe de trabajos prácticos en la cátedra de Fundaciones de la carrera de Ingeniería Civil de la Facultad de Ingeniería Civil de la Facultad de Ingeniería de la Universidad Nacional de Salta.
falbarracin@ucasal.edu.ar

ALDERETE, JOSÉ LUIS

Estudiante de Ingeniería Industrial en la Universidad Católica de Salta. Ayudante del Laboratorio de Materiales Avanzados del Grupo Interdisciplinario en Materiales GIM-IESIING de la UCasSal, con especialización en Materiales Magnéticos.
josealderete7@gmail.com

ARAMAYO, NORBERTO

Ingeniero en Informática y Técnico Universitario en Informática - Universidad Católica de Salta, Cursó Ingeniería Eléctrica Electrónica - Universidad Nacional. Córdoba Administrador de Base de Datos Oracle - Universidad Nacional De Jujuy – Universidad Oracle (EE.UU), HTML5 y CSS3 - Universidad Tecnológica Nacional, Diplomado en Moodle - Universidad Tecnológica Nacional – Capacitador en Informática - Estudios, desarrollo e implementación en Comunicación de Datos a Través de PLC (Power Line Communication) para el desarrollo de la transmisión de Internet por la red eléctrica, Desarrollo de software para el envío de mensajes de texto y Correo Electrónico, Transmisión de Streaming de audio y video- Desarrollo Web dinámicas – Asesor Tecnológico de los Laboratorios de Informática y Miembro del grupo de Robótica de la Facultad de Ingeniería de la UCASAL, Docente de la Cátedra de Sistemas Expertos de la Universidad Católica de Salta
saltaramayo@yahoo.com.ar

BOTELLI, MYRIAM CECILIA

Ingeniera Hidráulica, egresada de la Universidad Nacional de Salta. Especialista en Formulación, Evaluación de Proyectos de Inversión. Encargada del Gabinete de Hidráulica e instalaciones Sanitarias de la Universidad Católica de Salta. Docente de las cátedras de Hidrología, y de Hidráulica Aplicada en la carrera de Ingeniería Civil de la Universidad Católica de Salta.

gabinetehidraulica@ucasal.net

CATALINI, CARLOS GASTÓN

Ingeniero Civil egresado de la Universidad Católica de Córdoba. Profesor Titular de la cátedra de Hidráulica I de la carrera de Ingeniería Civil de la Universidad Católica de Córdoba. Co-director del Grupo de Investigación de Estudios Hidrológicos en Cuencas Pobrementemente Aforadas (EHCPA). A Cargo de la Dirección del Centro de Investigaciones de la Región Semiárida del Instituto Nacional del Agua (INA-CIRSA). Autor de numerosos artículos técnicos asociados a la temática hidrológica.

cgcatalini@hotmail.com

GAMARRA CARAMELLA, SOLEDAD

Ingeniera Química, egresada de la Universidad Nacional de Salta, donde actualmente cursa el Doctorado en Ingeniería. Posee una Beca Interna para Finalización del Doctorado otorgada por CONICET, con lugar de trabajo en la Universidad Católica de Salta donde desarrolla el trabajo de tesis titulado: «Materiales Magnéticos de Estructura Amorfa y Nanométrica». A partir del año 2013 se desempeña como Jefe de Trabajos Prácticos (dedicación simple) en la cátedra Termodinámica I, de la carrera Ingeniería Química de la Universidad Nacional de Salta.

gamarrasoledad@gmail.com

GARCÍA, CARLOS MARCELO

Ingeniero Civil egresado de la Universidad Católica de Córdoba. Magíster en Ciencias de la Ingeniería – Mención Recursos Hídricos egresado de la Universidad Nacional de Córdoba. Doctor: Dpto. de Ing. Civil y Ambiental, Universidad de Illinois en Urbana, Champaign, Estados Unidos. Investigador adjunto, Carrera del investigador científico y tecnológico del CONICET. Docente-Investigador, Facultad de Ciencias Exactas, Físicas y Naturales. Universidad Nacional de Córdoba (FCEFyN – UNC). Investigador Asociado del Instituto Nacional del Agua. Investigador visitante al United States Geological Survey, Office of Surface Water (Estados Unidos), USGS Illinois Water Science Center, Illinois, Estados Unidos. Director del Dpto. de Hidráulica de la Universidad Nacional de Córdoba. Director del CETA (Centro de Estudios y Tecnología del Agua) de la Universidad Nacional de Córdoba. Docente de cursos de grado y postgrado no solo en la Universidad Nacional de Córdoba sino también en diferentes instituciones educativas de la Argentina.

cgarcia2mjc@gmail.com

GARDEL, ANA EUGENIA ELVIRA

Ingeniera en Informática, egresada de la Universidad Católica de Salta. Se desempeña como Profesora Adscripta en la Cátedra de Tecnología Emergente I– Universidad Católica de Salta. En el Área de Investigación se especializa en programación para robótica. Participo como docente en los Talleres introducción a la programación de robots dictados en los años 2014 y 2015 en la Facultad de Ingeniería – Universidad Católica de Salta.
aee62@hotmail.com

GIUBERGIA, JORGE HORACIO:

Ingeniero Industrial de la Facultad de Ingeniería de la Universidad de Buenos Aires. Especialista en Aplicaciones Tecnológicas de la Energía Nuclear del Instituto Balseiro, Especialista en Reactores Nucleares, Dan Beninson. Responsable del Laboratorio de Física de la Ucasal y Director del Proyecto de investigación «Evaluación de la sustentabilidad del reemplazo de los transformadores eléctricos de distribución tradicionales por transformadores con núcleos amorfos en la provincia de Salta» Director General de Energía Eléctrica de la Secretaría de Energía de Salta.

GUILLÉN, NICOLÁS FEDERICO

Ingeniero Civil y Magíster en Ciencias de la Ingeniería – Mención Recursos Hídricos, egresado de la Universidad Nacional de Córdoba. En la actualidad es estudiante de doctorado y becario de CONICET. En el Área de Investigación se especializa en Hidrología e Hidráulica. En el año 2015 estuvo en el United States Geological Survey, Office of Surface Water (Estados Unidos) participando en el Water Science Center del estado de Illinois como investigador visitante.
nfguillen@hotmail.com

HAARALA MIKA

Es MSC en Matemáticas, por la Universidad de Joensuu, Finlandia. Es investigador del Instituto de Estudios interdisciplinarios de Ingeniería de la UCASAL, en el área de Matemática Aplicada y Estadística. Miembro del Grupo de riesgo Sísmico, con varias publicaciones sobre el tema.

MOYA, JAVIER

Ingeniero Mecánico con Orientación Diseño y Doctor en Ingeniería. Se desempeña como Profesor Extraordinario de Tecnologías y Procesos de Producción y de Conocimiento y Estudio de Materiales. Director del Laboratorio de Materiales Avanzados (FI-UCASAL) y Jefe del Departamento de Investigación de la Facultad de Ingeniería. Docente universitario (desde 1988) e Investigador del CONICET (desde 2005).
jmoya@ucasal.net

PÉREZ ABELLEIRA, ALICIA

Es Licenciada en Informática por la Universidad Politécnica de Madrid y PhD in Computer Science por Carnegie Mellon University. Actualmente se desempeña en la Facultad de Ingeniería de la Universidad Católica de Salta, como docente de Sistemas Expertos y de

Perfil académico-profesional de los autores

Compiladores y como Jefe del Departamento de Investigación. En la Universidad Carlos III de Madrid (España), fue docente de Inteligencia Artificial (2004, 2006) de la carrera de Ingeniería Informática- grupo bilingüe.
aperez@ucasal.net

OROSCO, LIA

Es Dra Ingeniera por la Universidad Politécnica de Cataluña, en el área de la Ingeniería Sísmica y Dinámica Estructural. Es Profesora Titular de la Cátedra Construcciones de Hormigón Armado, en la Facultad de Ingeniería de la UCASAL. Directora del Instituto de Estudios Interdisciplinarios de Ingeniería, y dirige el Grupo Riesgo Sísmico.
lorosco@ucasal.edu.ar>

SAID, CARLOS

Magister en Ingeniería del SW (I.T.B.A.), Magister en Ingeniería del Conocimiento (U.P.M.), Magister en Ciencia, Tecnología y Sociedad (Universidad de Quilmes), Postgrado en Dirección de Sistemas (Universidad de Belgrano), Licenciatura en Sistemas (Proc. Electrónico). Docente de grado, postgrado y maestría en diversas universidades públicas y privadas). Ex-decano de la Facultad de Ingeniería (UCASAL), Ex-coordinador de grupos de trabajo en el IRAM (Seguridad de la Inf., Calidad del SW, Riesgos). Publicaciones en medios escritos, radio, tv, Internet.
cgsaid@ucasal.edu.ar

TALAME, MARIA LORENA

Licenciada en Análisis de Sistemas, egresada de la Universidad Nacional de Salta. Actualmente, cursa el Master de Ingeniería Informática en la Universidad Abierta de Cataluña. Se desempeña como Profesora Adjunta en la Cátedra de Sistemas Expertos y en la Cátedra de Tecnologías Emergentes I de la carrera de Ingeniería en Informática de la Universidad Católica de Salta.
lorenatalame@gmail.com

ZAMBRANO ECHENIQUE, MANUEL LUIS

Ingeniero Industrial, especialista en Higiene y Seguridad en el trabajo, egresado de la Universidad Católica de Salta. Se desempeña como coordinador de la UT Innovación Tecnológica y Proyectos Tecnológicos en INTI SALTA, Coordinador de Laboratorios de la UCASAL y Profesor en distintas cátedras (Organización Industrial; Logística; Higiene y Seguridad e Higiene y Seguridad Industrial) de la carrera de Ingeniería Civil- Industrial-Informática y Telecomunicaciones de la Universidad Católica de Salta. En el sector productivo local dirige proyectos PROCODAS - FONARSEC y PFIP.
mzambrano@ucasal.edu.ar

UNIVERSIDAD CATÓLICA DE SALTA

AUTORIDADES

Rector	Ing. Rodolfo Gallo Cornejo
Vicerrectora Académica	Mg. Constanza Diedrich
Vicerrector de Investigación y Desarrollo	Dr. Federico Colombo Speroni
Vicerrector Administrativa	Cr. Roberto Dib Ashur
Vicerrector de Fomación	Pbro. Dr. Cristian Arnaldo Gallardo
Director General del Sistema de Educación a Distancia	Ing. Lic. Daniel Torres Jiménez
Secretaria General	Lic. Silvia Milagro Álvarez

AUTORIDADES DE LA FACULTAD DE INGENIERÍA

Decano de la Facultad de Ingeniería	Mg. Ing. Néstor Eugenio Lesser
Secretario Académico	Ing. Mario Alfredo Ochoa
Secretaria Técnica	Sec. Ejec. Andrea Avelina Burgos
Jefa del Departamento de Investigación	Dra. Ing. Lía Elizabet Orosco Segura
Jefe del Departamento de Extensión, Graduados y Bienestar	Ing. Juan Francisco Linares
Jefe del Departamento de Ingeniería Industrial	Esp. Ing. Manuel Luis Zambrano Echeñique
Jefe de Departamento de Ingeniería Civil	Ing. Fernando Javier Albarracín
Jefe del Departamento de Ingeniería en Informática	Ing. Guillermina Nievas
Jefe del Departamento de Ingeniería en Telecomunicaciones	Ing. Roberto Daniel Breslin
Coordinador de Laboratorios	Esp. Ing. Mirian Cecilia Botelli

EUCASA

Ediciones Universidad Católica de Salta

Directora	Lic. Rosanna Caramella
Edición	Prof. Soledad Martínez
Comercialización	Lic. Mariana Remaggi