



Universidad Veracruzana

UNIVERSIDAD VERACRUZANA

FACULTAD DE ESTADÍSTICA E
INFORMÁTICA

Pronóstico de la Demanda en los Sistemas de
Distribución de Energía Eléctrica: Aplicando
la Teoría de Valores Extremos

Tesis

Que como requisito parcial para obtener el grado de

MAESTRO EN ESTADÍSTICA

Presenta:

Carlos González Andrade

TUTOR

Dr. Sergio Francisco Juárez Cerrillo

Xalapa, Enríquez, Ver. Julio de 2010

Dedicatoria

Agradecimientos

Resumen.

El diseño técnico-económico de los sistemas de distribución de energía eléctrica requiere fundamentarse en pronósticos precisos de la demanda máxima, y en la evaluación objetiva de los riesgos financieros asociados con diferentes alternativas de diseño. La Teoría de Valores Extremos y los principios asociados al desarrollo de pequeñas áreas se aplican en este trabajo para diseñar un método de pronóstico que pueda usarse con los registros mensuales de la demanda máxima, disponibles en las subestaciones de la CFE, y que permita obtener una medida del riesgo asociado con la selección de la capacidad requerida para el diseño de este tipo de sistemas.

Esta metodología se ilustra con los datos de una subestación de distribución, ajustando dos modelos basados en la Distribución de Valor Extremo Generalizado (DVEG). Los modelos incorporan dos curvas de crecimiento distintas, asociadas con el parámetro de localización, para describir la tendencia de crecimiento de la demanda a largo plazo. Los cuantiles de la DVEG de estos modelos se utilizan para pronosticar la demanda máxima del área atendida por la subestación en función de la probabilidad de exceder el valor pronosticado. El horizonte de pronóstico se extiende 20 años.

TABLA DE CONTENIDO.

DEDICATORIA	i
AGRADECIMIENTOS.....	ii
TABLA DE CONTENIDO	iv
LISTA DE FIGURAS.	vii
LISTA DE TABLAS.	viii
CAPÍTULO 1. INTRODUCCIÓN.....	1
1.1 Pronóstico de la demanda.	2
1.1.1 Horizontes de pronóstico.....	3
1.1.2 Técnicas de pronóstico.....	4
1.1.3 Pronósticos de áreas pequeñas.....	5
1.2 Teoría de valores extremos.	7
1.3 Problema.....	8
1.4 Objetivos.....	9
1.5 Organización de la investigación.	10
CAPÍTULO 2. MÉTODOS DE TENDENCIA.....	11
2.1 Métodos deterministas.....	13
2.1.1 Crecimiento lineal.	14
2.1.2 Crecimiento exponencial.	14
2.1.3 Otras curvas de crecimiento.....	15
2.1.3.1 Crecimiento exponencial modificado.....	15
2.1.3.2 Curva logística simple.	16

2.1.3.3	Curva S de Brown, Hanson y Hagan.	17
2.1.3.4	Polinomio de Willis.	18
2.2	Modelos estadísticos:.....	19
CAPÍTULO 3. MODELOS DE VALORES EXTREMOS.		22
3.1	Distribución del Valor Extremo Generalizado.	23
3.2	Teorema de Fisher-Tippet.....	24
3.3	Inferencia en la DVEG.	26
3.3.1	Estimación de parámetros de la DVEG.	27
3.3.2	Estimadores de máxima verosimilitud.	28
3.3.3	Error estándar de los estimadores.....	29
3.4	Cuantiles de la DVEG.	29
3.5	Modelos que incorporan tendencia.	31
3.6	Diagnóstico del modelo.....	34
CAPÍTULO 4. CASO DE ESTUDIO.		35
4.1	Datos de la subestación.....	36
4.2	Modelos para el parámetro de localización.....	37
4.2.1	Modelo lineal.....	37
4.2.2	Modelo logístico.	38
4.3	Modelo de la DVEG con tendencia lineal en el parámetro de localización.....	38
4.3.1	Estimación de los parámetros de la DVEG.	40
4.3.2	Diagnóstico del modelo.....	42
4.3.3	Cuantiles de la distribución de la DMA.....	42
4.3.4	Errores estándar e intervalos de confianza de los cuantiles de la DMM.	46
4.3.5	Pronóstico de la DMA de la subestación.....	50
4.4	Modelo de la DVEG con tendencia logística en el parámetro de localización.....	52
4.4.1	Parámetros estimados de la DVEG.	53
4.4.2	Diagnóstico del modelo.....	55

4.4.3	Cuantiles de la distribución de la DMA.....	56
4.4.4	Cálculo de los errores estándar.....	57
4.4.5	Pronóstico de la DMA.....	59
CAPÍTULO 5. CONCLUSIONES Y RECOMENDACIONES.....		61
5.1	Conclusiones.....	61
5.2	Ventajas y desventajas.....	63
5.3	Contribución.....	64
5.4	Investigación actual y futura.....	65
Referencias.....		67

Lista de figuras.

Figura 1-1. Demanda máxima mensual en una subestación de distribución.	9
Figura 2-1. Registro mensual de la demanda máxima de un sistema de distribución, (unidades en MW).	12
Figura 2-2. Comportamiento a largo plazo de las distintas curvas de tendencia.	19
Figura 3-1. Gráfico de niveles de retorno.	31
Figura 4-1. DMM en la subestación. La línea continua corresponde al período de invierno y la línea punteada corresponde al período de verano.	37
Figura 4-2. Diagnóstico del ajuste del modelo de la DVEG con tendencia lineal.	42
Figura 4-3. Cuantil 0.5 de la distribución estimada de la DMA para una dinámica de crecimiento lineal sin límite.	43
Figura 4-4. Pronóstico de la DMA, para la temporada de verano de los años 3, 5, 10, 15 y 20 futuros respecto a la probabilidad de exceder el valor pronosticado.	44
Figura 4-5. Pronóstico de la DMA, para la temporada de invierno de los años 5, 10, 15 y 20 futuros a la probabilidad de exceder el valor pronosticado.	46
Figura 4-6. Gráfica del error estándar de los cuantiles, con un horizonte de pronóstico a 20 años, usando el modelo de tendencia lineal.	48
Figura 4-7. Pronóstico a 20 años de la DMA de la subestación, con una probabilidad de 0.1%, 1%, 10% y 50% de exceder el valor pronosticado.	51
Figura 4-8. Diagnóstico del ajuste del modelo de la DVEG con tendencia logística.	55
Figura 4-9. Pronóstico de la DMA, para los años 3, 5, 7, 10, 15 y 20 futuros, en función de la probabilidad de exceder el valor pronosticado.	56
Figura 4-10. Gráfica del error estándar de los niveles de retorno para los años de pronóstico 3, 5, 10, 15 y 20, usando el modelo de tendencia logístico.	58
Figura 4-11. Pronóstico a 20 años de la DMA de la subestación, considerando una probabilidad del 2.5% de exceder el valor pronosticado.	59

Lista de tablas.

Tabla 3-1. Muestra de máximos basada en máximos muestrales de m períodos de tiempo	27
Tabla 4-1. DMM para la temporada de invierno.	36
Tabla 4-2. DMM para la temporada de verano.	36
Tabla 4-3. Estimaciones de máxima verosimilitud de los parámetros del modelo, (errores estándares en paréntesis).....	40
Tabla 4-4. Matriz aproximada de varianzas y covarianzas de los parámetros estimados de la DVEG.	40
Tabla 4-5. Gradiente asociado al cuantil $(1 - p)$ de la DVEG.	47
Tabla 4-6. DMA pronosticada en la subestación, (temporada de invierno).	49
Tabla 4-7. DMA pronosticada en la subestación, (temporada de verano).	49
Tabla 4-8. Estimaciones de máxima verosimilitud de los parámetros del modelo, (errores estándares en paréntesis).....	54
Tabla 4-9. DMA pronosticada en la subestación, (tendencia logística).	58

CAPÍTULO 1. INTRODUCCIÓN.

Los sistemas de distribución de energía eléctrica, SDEE, deben satisfacer la demanda de energía eléctrica de la sociedad de forma segura y económica, y con un grado aceptable de calidad y confiabilidad. En general, estos sistemas se diseñan de forma robusta para satisfacer la demanda máxima bajo condiciones operativas extremas (Willis, 2004). Sin embargo, para satisfacer el crecimiento de la demanda, originado en la incorporación de nuevos servicios y el aumento en el consumo de los servicios existentes (Willis, 2002), debe invertirse continuamente en el desarrollo de su infraestructura.

El diseño económico y técnicamente adecuado de las obras de expansión de un SDEE, requiere del conocimiento anticipado y preciso de los niveles potenciales máximos de la demanda en diferentes partes del

sistema (Huck et al., 1980), así como la evaluación de los riesgos financieros asociados con la incertidumbre en el pronóstico de estos niveles potenciales (Hyndman & Fan, 2008).

1.1 Pronóstico de la demanda.

Los pronósticos de demanda máxima se emiten para guiar decisiones durante la planificación de un sistema eléctrico, éstas, de acuerdo con Diebold (1999), implican cierto grado de pérdida que se incrementa a medida que la decisión elegida se aleja más de la decisión correcta (aquella cuya pérdida es cero).

Por ejemplo, si la demanda máxima pronosticada para un año en particular excede la capacidad del sistema, será necesario invertir en la construcción de algunas obras de infraestructura que eviten su saturación en la fecha pronosticada. Tal decisión lleva asociado el riesgo de que la demanda máxima pueda exceder esta capacidad antes de lo pronosticado, dando lugar a pérdidas económicas y sociales por la energía no suministrada y la insatisfacción de los usuarios.

Además, de acuerdo con Engel y Palmer (1992), la planificación de un SDEE debe lograr la mayor economía en la operación de un SDEE,

luego entonces, se requiere minimizar los excedentes de capacidad para diferentes condiciones operativas, evitando anticipar innecesariamente los proyectos de inversión, que además de su baja rentabilidad económica implicarán altos costos operativos, con repercusiones en la estructura tarifaria y en la economía de la sociedad.

1.1.1 Horizontes de pronóstico.

La industria eléctrica requiere pronósticos de energía y demanda a corto, mediano y largo plazo, con horizontes que varían de 1 a 7 días, de 1 a 5 años, y de 5 a 30 años, respectivamente (Elias & Hatziargyriou, 2009). Y dado que el tiempo requerido para construir las principales obras de expansión en los SDEE varía de 3 a 5 años, los pronósticos a mediano y largo plazo son de particular interés en la planificación de este tipo de sistemas.

Por otra parte, con base en pronósticos a largo plazo se establecen planes estratégicos, se juzgan nuevas ideas, se analizan cambios a la normatividad (Engel & Palmer, 1992), y se anticipan nuevas obras de infraestructura, permitiendo su ajuste a cambios imprevistos originados por incertidumbres en la demanda, la disponibilidad de sitios, demoras en

la entrega de equipos, costos de equipo y combustibles, cambios tecnológicos, políticas de operación, requerimientos de confiabilidad, etc. (Henault et al., 1970).

1.1.2 Técnicas de pronóstico.

Desde los años 70's existe gran variedad de técnicas para el pronóstico de la demanda en sistemas eléctricos (Gellings, 1996). Recientemente, la investigación se centra en técnicas de pronóstico a corto y muy corto plazo (Lotufo & Minussi, 1999), las cuales, de acuerdo con Alfares y Nazeeruddin (2001), pueden clasificarse en las siguientes categorías: (1) regresión múltiple, (2) suavizamiento exponencial, (3) mínimos cuadrados ponderados iterativamente, (4) pronóstico de carga adaptativo, (5) series estocásticas de tiempo, (6) modelos ARMAX basados en algoritmos genéticos, (7) lógica difusa, (8) redes neuronales, y (9) sistemas expertos basados en el conocimiento.

Para el pronóstico a mediano y largo plazo se aplican modelos econométricos, métodos de uso final, y de tendencia (Genethliou, 2005), para los cuales el análisis multiescenario se considera como la única forma práctica de tratar con la incertidumbre (Willis, 2002).

Los modelos econométricos combinan las técnicas estadísticas y la teoría económica estableciendo una relación causa-efecto entre la demanda y variables económicas y demográficas (Fu & Nguyen, 2003). Los modelos de uso final pretenden explicar el comportamiento de la demanda con relación al número de aparatos eléctricos utilizados en los servicios residenciales, comerciales e industriales (Gellings, 1996). Mientras que los métodos de tendencias se basan en la extrapolación de curvas de crecimiento ajustadas a datos históricos (Kandil et al., 2001).

1.1.3 Pronósticos de áreas pequeñas.

Un SDEE se compone de gran número de unidades dispersas en un área geográfica, cuya demanda no se distribuye de manera uniforme (Willis, 2002). Por tanto se requieren pronósticos con un nivel de granularidad suficiente para estimar la demanda en cada una de sus subestaciones y circuitos de distribución.

Con este fin, en los años 70's, se diseñaron las técnicas para el pronóstico de "áreas pequeñas" (Menge et al., 1977), que consisten en modelar su dinámica de desarrollo hasta obtener un valor de saturación asintótico, considerando que su superficie se mantiene constante. Estas

técnicas utilizan principalmente métodos de simulación para modelar la interacción de la demanda con factores urbanos y demográficos (Willis y Northcote-Green, 1982).

Sin embargo, la complejidad de los métodos de simulación (Temraz & Quintana, 1992, y Brown et al., 1999), y la falta de información suficiente para aplicar métodos econométricos y de uso final, hace que muchas empresas eléctricas prefieran, debido a su sencillez, el uso de métodos determinísticos para extrapolar la demanda máxima de subestaciones y circuitos de distribución (Gellings, 1996).

No obstante, la demanda máxima de los sistemas eléctricos está sujeta a un gran rango de incertidumbres (Henault et al., 1970, Hyndman & Fan, 2008), que contradicen las implicaciones (Nelson, 1973) de los modelos determinísticos sobre su evolución completamente sistemática y altamente predecible. Por tanto, se requieren métodos estadísticos que permitan evaluar de manera objetiva el riesgo asociado con el proceso de toma de decisiones (Montgomery, 1990).

1.2 Teoría de valores extremos.

La aplicación de métodos estadísticos al análisis de la demanda máxima de los SDEE depende, fundamentalmente, de comprender la distribución de probabilidad que gobierna sus valores máximos potenciales; los cuales, por definición, corresponden a uno de los valores extremos del conjunto de valores que puede tomar la demanda de un sistema eléctrico durante cierto período de tiempo (diario, mensual, anual, etc.) (Belzer y Kellogg, 1993).

La Teoría de Valores Extremos (TVE) es una de las ramas de la probabilidad y estadística que proporciona las bases para hacer inferencias sobre la dimensión y frecuencia de este tipo de eventos, cuyo comportamiento depende del extremo de una distribución de probabilidad, por lo que resulta muy adecuada para analizar el comportamiento de la demanda máxima, en contraste con las técnicas de regresión y los modelos de series de tiempo, que se enfocan, en el centro de dicha distribución.

La TVE tiene aplicaciones importantes en el área de riesgo y cuenta con extensas aplicaciones en ingeniería, hidrología, economía y meteorología (Castillo, 1988). Algunas de sus aplicaciones al pronóstico de

la demanda máxima en sistemas eléctricos se encuentran en Belzer y Kellogg (1993), y en Loredó (2001), en estos trabajos se analizan algunas de las fuentes de incertidumbre que intervienen en este tipo de pronósticos, y se obtienen medidas de su grado de precisión.

1.3 Problema.

En este trabajo se aborda el problema de diseñar un procedimiento de pronóstico fundamentado en la TVE para estimar la demanda máxima de energía eléctrica en sistemas de distribución de energía eléctrica, con horizontes a mediano y largo plazo, considerando los datos disponibles en la Comisión Federal de Electricidad (CFE), que consisten en un conjunto de series de tiempo con el registro mensual de la demanda máxima en cada una de sus subestaciones (Figura 1-1), sin información adicional que permita correlacionar la demanda con variables climatológicas o económicas.

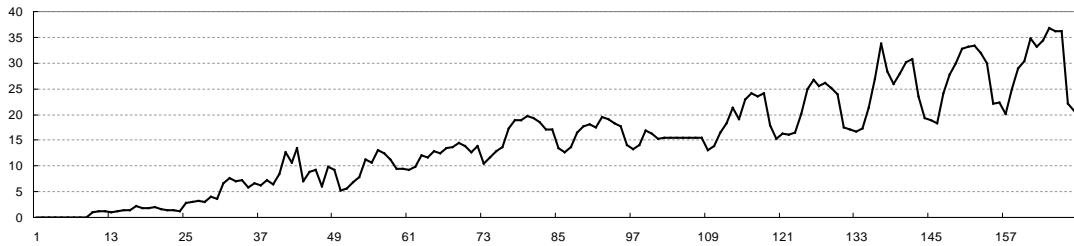


Figura 1-1. Demanda máxima mensual en una subestación de distribución.

Los pronósticos que se obtienen actualmente a partir de estos datos

se basan en técnicas de extrapolación determinísticas que descartan gran parte de los datos disponibles, al considerar únicamente los valores anuales máximos y no consideran el riesgo inherente a este proceso.

1.4 Objetivos.

El objetivo general de este trabajo es el de investigar la aplicación de la TVE para pronosticar la demanda máxima anual de energía eléctrica en SDEE, a fin de obtener una medida del riesgo asociado con la selección de su capacidad de diseño, considerando las diferentes alternativas que se evalúan durante el proceso de planificación.

Particularmente se propone una metodología para pronosticar la demanda máxima anual en subestaciones de distribución que aproveche el mayor número de los datos disponibles en la CFE, e incorporar los

conceptos sobre el desarrollo de pequeñas áreas para modelar la dinámica de crecimiento de la demanda en este tipo de sistemas.

1.5 Organización de la investigación.

El resto del trabajo está organizado de la siguiente manera. En el Capítulo 2, se describen las principales técnicas de extrapolación determinísticas, se destacan los modelos utilizados para describir la dinámica de crecimiento de la demanda en pequeñas áreas. Los principios para la modelación de valores extremos, a través de la llamada distribución del valor extremo generalizada, se introducen en el Capítulo 3. En el Capítulo 4 se muestra la aplicación de la TVE al pronóstico de la demanda máxima mediante el análisis de los datos de demanda máxima mensual registrados en una subestación de distribución de CFE; y se muestra la forma de incorporar los métodos tendencias en los modelos de valores extremos, para describir la dinámica de crecimiento en esta subestación usando los principios del desarrollo de áreas pequeñas. En el Capítulo 5 se discuten las ventajas y desventajas de la propuesta y se indican algunas líneas de investigación para mejorar el modelo propuesto.

CAPÍTULO 2. MÉTODOS DE TENDENCIA.

La demanda de energía eléctrica de un sistema de distribución representa la tasa de transferencia de energía que se requiere a cada momento para hacer funcionar el equipamiento eléctrico de una población. Su comportamiento se ve afectado por el crecimiento demográfico y el grado de desarrollo económico (Murto, 1998, Mirasgedis et al., 2006), así como por distintos fenómenos naturales y sociales (Kermanshahi, 1993).

Los SDEE que interactúan muy poco con otros sistemas, observan patrones de comportamiento de su demanda máxima mensual, DMM, como el que se muestra en la Figura 2-1, en la cual se evidencia cierta variación estacional y una clara tendencia en el crecimiento a largo plazo.

Para describir la dinámica de crecimiento observada en este tipo de sistemas, frecuentemente se utilizan técnicas de extrapolación

determinísticas, que involucran el ajuste de curvas de crecimiento a los datos históricos, con las cuales se obtienen pronósticos al extrapolar sus valores futuros (Willis, 2002), consiguiendo, en algunos casos, resultados razonables.

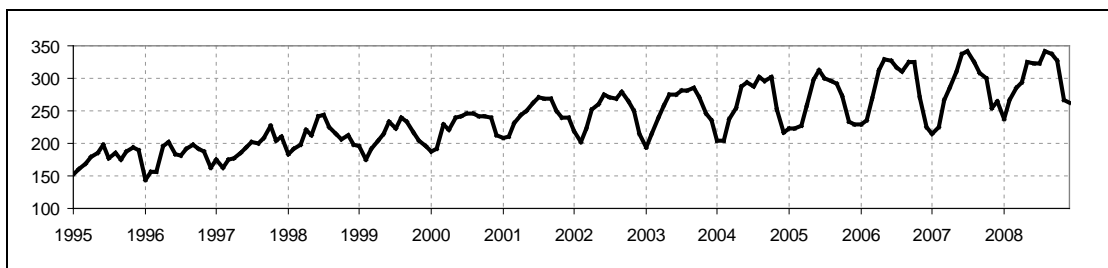


Figura 2-1. Registro mensual de la demanda máxima de un sistema de distribución, (unidades en MW).

Sin embargo, la regularidad exhibida por algunos datos históricos puede crear la falsa impresión respecto a la precisión de estos métodos. Debe considerarse que la proyección de tendencias no es un pronóstico en sí mismo, sino una herramienta útil para su análisis exploratorio (Willis, 1979).

La incertidumbre de los factores que influyen en el crecimiento de la demanda requiere de procedimientos de pronóstico que permitan tomar en cuenta el riesgo asociado de una manera objetiva, a través de la

varianza de los estimadores, y la obtención de intervalos de predicción probabilísticos (Henault et al., 1970).

A continuación se describen algunas de las curvas de crecimiento frecuentemente utilizadas para pronosticar la demanda máxima en los sistemas de distribución.

2.1 Métodos deterministas.

Los métodos de tendencias se basan en el ajuste, por mínimos cuadrados, de una curva de crecimiento a los datos históricos, cuya forma se selecciona *a priori* con base en argumentos teóricos. Esta relación funcional se utiliza para extrapolar la demanda en forma determinista, asumiendo que las condiciones que originaron su dinámica pasada serán las mismas en el futuro (Stanton & Gupta, 1970).

Ya que el crecimiento implica una tasa de cambio, estos modelos involucran ecuaciones diferenciales. Con este enfoque, el comportamiento de la demanda máxima se describe por medio de una función determinística del tiempo, $x(t)$, y su tasa de crecimiento por su derivada

$x'(t)$, cuyas propiedades determinan la función $x(t)$, inicialmente desconocida (Gordon, 1978).

2.1.1 Crecimiento lineal.

El modelo más sencillo se obtiene considerando una tasa de crecimiento constante. Así, $x'(t) = k$, implica que $x(t) = x_0 + kt$, para un valor inicial, $x_0 = x(0)$. Aplicando el método de mínimos cuadrados, esta línea de tendencia se ajusta a los datos históricos para determinar x_0 y k . El pronóstico se obtiene evaluando $x(t)$ para valores de t futuros.

2.1.2 Crecimiento exponencial.

Para Gönner (1986), el pronóstico de la demanda en SDEE puede obtenerse aplicando modelos de crecimiento exponencial. En éstos, la *tasa relativa de crecimiento*, definida por: $x'(t)/x(t)$ (Grossman, 1996), se mantiene constante. Su ecuación diferencial: $x'(t) = k x(t)$, tiene como solución: $x(t) = x_0 \exp(kt)$, para un valor inicial, $x_0 = x(0)$. Aplicando el método de mínimos cuadrados, esta curva de crecimiento se ajusta a los datos históricos para determinar los parámetros x_0 y k mediante la transformación siguiente: $\ln x(t) = \ln x_0 + kt$.

2.1.3 Otras curvas de crecimiento.

Considerando que los modelos lineal y exponencial implican el crecimiento indefinido de la demanda, Willis (1984) y Brown et al. (1999), recomiendan el uso de curvas de crecimiento sigmoides para pronosticar la demanda en áreas pequeñas, asumiendo que, una vez que la superficie se halla completamente desarrollada, su demanda se aproximará a un valor asintótico.

Entre estas curvas de crecimiento se encuentran las funciones: exponencial modificada, logística, y de Gompertz.

2.1.3.1 Crecimiento exponencial modificado.

En este modelo se considera que el crecimiento de la demanda en un área pequeña se encuentra limitado al valor asintótico, X , y que su tasa de crecimiento, en un momento dado, es proporcional a la demanda aún no satisfecha. Si $x(t)$ representa la demanda de la superficie ocupada al tiempo t , el modelo plantea la ecuación diferencial $x'(t) = k(X - x(t))$, cuya solución es $x(t) = X(1 - \exp(-k t)) + x_0 \exp(-k t)$, para una demanda

inicial, $x_0 = x(0)$. Cuando esta última es nula, se tiene el caso particular $x(t) = X(1 - \exp(-k t))$.

2.1.3.2 Curva logística simple.

En este modelo se considera que la tasa de crecimiento relativa es proporcional a la demanda aún no satisfecha. Su ecuación diferencial $x'(t)/x(t) = k(X - x(t))$, tiene la solución $x(t) = X\{1 + A \exp(-kXt)\}^{-1}$, donde A es una constante de integración que depende del momento inicial t_0 (Bossert, 1977).

Aplicando el método de mínimos cuadrados y la selección heurística del valor asintótico razonable, X , se determinan los parámetros A y k considerando la siguiente transformación (Willis, 1979)

$$\ln\left(\frac{X}{x(t)} - 1\right) = \ln A - kXt.$$

En este modelo la tasa de crecimiento inicia con un valor relativamente bajo, creciendo inicialmente como lo haría el modelo exponencial. Eventualmente la tasa de crecimiento comienza a declinar, continuando como lo haría en el modelo exponencial modificado, hasta

volverse finalmente asintótica en el valor de saturación esperado, adoptando la forma sigmoide que lo caracteriza.

Alternativamente al modelo logístico, se utiliza frecuentemente el modelo de Gompertz, cuya ecuación es $x(t) = Xa^{g^t}$. Schmill (1968) utilizó este modelo en 1968 para pronosticar la demanda de energía eléctrica en México hacia el año 2000.

2.1.3.3 Curva S de Brown, Hanson y Hagan.

Brown et al. (1999) proponen una relación para la tasa de crecimiento de la demanda en un área pequeña que es directamente proporcional tanto a la demanda aún no satisfecha, como al tiempo transcurrido desde que el área inició su desarrollo, $(t - t_0)$. La ecuación diferencial resultante es $x'(t) = k(t - t_0)(X - x(t))$, y su solución tiene la forma siguiente (haciendo $k = 2/\tau^2$):

$$x(t) = X \left(1 - \exp \left\{ - \left(\frac{t - t_0}{\tau} \right)^2 \right\} \right); \text{ para } t > t_0,$$

donde τ es una constante de tiempo cuyo valor se determina de manera heurística, al igual que el valor asintótico, X , y la fecha de inicio del desarrollo t_0 , (Brown et al., 1999).

2.1.3.4 Polinomio de Willis.

Willis (2002) recomienda una función polinomial de tercer grado, $x(t) = b_0 + b_1t + b_2t^2 + b_3t^3$, para modelar el crecimiento de la demanda en áreas pequeñas. Aplicando el método de mínimos cuadrados, este polinomio se ajusta a los datos históricos y a un valor de saturación, X , el cual se determina heurísticamente, para uno o más años horizonte, con el fin de que el modelo asemeje la forma de una curva de crecimiento sigmoide.

La Figura 2-2 compara el comportamiento de las diferentes curvas de tendencia ajustadas a los datos de la Figura 2-1. Se observa que, aunque el polinomio cúbico requiere del año horizonte para mejorar el pronóstico, el comportamiento a largo plazo queda determinado por su componente de mayor grado, por lo que su extrapolación carece de valor práctico más allá del valor horizonte establecido.

Por otra parte, se aprecia que las curvas de crecimiento logística, y de “S de Brown”, modelan de forma adecuada el comportamiento a largo plazo para el valor asintótico seleccionado. En el siguiente capítulo se muestra la forma de incorporar estos modelos para el pronóstico de la demanda máxima utilizando la Teoría de Valor Extremo.

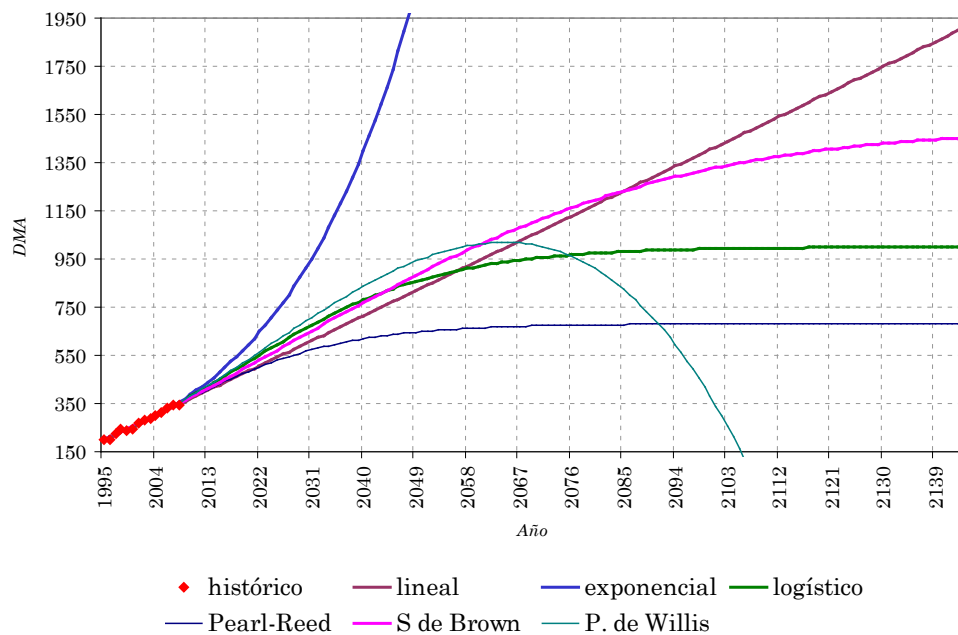


Figura 2-2. Comportamiento a largo plazo de las distintas curvas de tendencia.

2.2 Modelos estadísticos:

De acuerdo a Nelson (1973), cualquier modelo que represente una variable como función determinística del tiempo implica su evolución a

largo plazo completamente sistemática, y altamente predecible. Por tanto, con el fin de considerar la incertidumbre en los modelos de pronósticos se requiere incluir en los modelos un elemento aleatorio, $\varepsilon(t)$, que permita tomar en cuenta los errores de pronóstico.

De este modo, si el valor esperado de la demanda se representa por la función determinística del tiempo, $f(t)$, su comportamiento estadístico a través del tiempo se representa por

$$X(t) = f(t) + \varepsilon(t).$$

Las propiedades estadísticas de este modelo dependen de la distribución de probabilidad asociada al error. En consecuencia, $X(t)$ es una variable aleatoria para cada valor de t , que puede estudiarse con técnicas de regresión y series de tiempo, cuya aplicación se basa generalmente en los siguientes supuestos (Willis, 1979):

- La variable independiente se conoce con precisión.
- La relación a ser ajustada es lineal en los parámetros y en los residuos $\varepsilon(t)$.

- El error es una variable aleatoria normalmente distribuida con media 0 y varianza constante y además $\text{Cov}(\varepsilon(s), \varepsilon(t)) = 0$, para $s \neq t$.

Los modelos de regresión se utilizan ampliamente (Montgomery, 1990) para predecir los valores futuros de la demanda, siendo el valor esperado $E[X(t)] = f(t)$ igual a su estimación puntual. De acuerdo con Stanton (1969) dicha estimación implica un riesgo del 50% de exceder el valor pronosticado.

CAPÍTULO 3. MODELOS DE VALORES EXTREMOS.

La distribución de valor extremo identificada por Fisher y Tippett (1928) y desarrollada posteriormente por Gumbel (1958), Galambos (1978) y Leadbetter (1983), entre otros, se utiliza para hacer inferencias con respecto a los máximos o mínimos de un conjunto de datos.

En el desarrollo clásico de la TVE, los valores extremos corresponden a los máximos o mínimos de un conjunto de muestras. Por tanto, el análisis estadístico de la demanda máxima de energía eléctrica equivale al análisis de la distribución de sus valores extremos (Belzer y Kellogg, 1993).

En este capítulo se describen de manera general los procedimientos de inferencia de la TVE que se utilizan en el siguiente capítulo para analizar la demanda máxima de energía eléctrica.

3.1 Distribución del Valor Extremo Generalizado.

Supongamos que Y_1, \dots, Y_n es una muestra aleatoria de cierta distribución desconocida, es decir de $F(y) = P(Y \leq y)$. Sea $X_n = \max\{Y_1, \dots, Y_n\}$ el máximo muestral. La TVE establece que, bajo condiciones muy generales, la distribución del máximo muestral X_n , definida por

$$P(X_n \leq x) = P(Y_1 \leq x, Y_2 \leq x, \dots, Y_n \leq x) = F^n(x), \quad (3-1)$$

se puede aproximar con la *Distribución del Valor Extremo Generalizada* (DVEG) cuya función de distribución está dada por

$$G(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad (3-2)$$

y cuya función de densidad de probabilidad es igual a

$$g(x; \mu, \sigma, \xi) = \frac{1}{\sigma} \left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-(1+1/\xi)} \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \quad (3-3)$$

para $\{x : 1 + \xi(x - \mu)/\sigma > 0\}$, $-\infty < \mu < \infty$, $\sigma > 0$ y $-\infty < \xi < \infty$ con $\xi \neq 0$; donde μ es un parámetro de localización, σ de escala y ξ de forma.

Cuando $\xi = 0$ la DVEG toma la forma de la distribución Gumbel:

$$G(x; \mu, \sigma) = \exp\left[-\exp\left\{-\left(\frac{x - \mu}{\sigma}\right)\right\}\right], \quad -\infty < x < \infty.$$

3.2 Teorema de Fisher-Tippet.

Los detalles matemáticos de la justificación para aproximar la distribución del máximo muestral con la DVEG, es decir, de $P(X_n \leq x) \approx G(x; \mu, \sigma, \xi)$ puede verse en los tratados de la TVE, por ejemplo en Coles (2001), Embrechts et al. (1997), y Reiss et al. (1997). A continuación se resume brevemente algunos aspectos de esta justificación.

La función $F^n(x)$ de la ec. (3-1) resulta en sí misma de poco interés, ya que se puede demostrar que para cualquier $x < x^*$, donde $x^* = \sup\{x \in \mathbb{R} : F(x) < 1\}$, se tiene que $P(X_n \leq x) \rightarrow 0$, conforme $n \rightarrow \infty$, mientras que $X_n \rightarrow x^*$ casi seguramente. Sin embargo, bajo ciertas circunstancias, se demuestra (véase Haan y Ferreira, 2006) que existen secuencias de constantes normalizadoras $a_n > 0$ y b_n , tales que:

$$P\left\{\frac{X_n - b_n}{a_n} \leq x\right\} = F(a_n x + b_n)^n \rightarrow G(x), \text{ conforme } n \rightarrow \infty \quad (3-4)$$

El Teorema de Fisher-Tippett (1928), véase también Gnedenko (1943), establece que si G es una función no degenerada, ésta pertenece a alguna de las siguientes familias de distribuciones:

$$I. \quad G(x) = \exp\left\{-\exp\left[-\left(\frac{x-b}{a}\right)\right]\right\}, -\infty < x < \infty \quad (\text{Gumbel})$$

$$II. \quad G(x) = \begin{cases} 0, & x \leq b \\ \exp\left\{-\left(\frac{x-b}{a}\right)^{-\alpha}\right\}, & x > b; \quad \alpha > 0 \end{cases} \quad (\text{Fréchet})$$

$$III. \quad G(x) = \begin{cases} \exp\left\{-\left(\frac{x-b}{a}\right)^\alpha\right\}, & x < b; \quad \alpha > 0 \\ 1, & x \geq b \end{cases} \quad (\text{Weibull})$$

Es decir, el máximo muestral adecuadamente normalizado $(X_n - b_n)/a_n$, converge en distribución a una variable aleatoria cuya distribución pertenece a alguna de las familias *I*, *II*, o *III*, denominadas colectivamente como *distribuciones de valor extremo*. La DVEG introducida en (3-2) es conocida también como la representación de Jenkinson-von Mises de las distribuciones del valor extremo. Cuando $\xi = 0$, en el sentido de que $\xi \rightarrow 0$, la DVEG se reduce a la distribución Gumbel. Y para $\xi > 0$ y $\xi < 0$ la DVEG se reduce a la Fréchet y Weibull, respectivamente.

De la ec. (3-4) tenemos que

$$P\{(X_n - b_n)/a_n \leq (x - b_n)/a_n\} \approx G((x - b_n)/a_n)$$

$$P\{X_n \leq x\} \approx G^*(x),$$

donde G^* es la DVEG. Por lo tanto, si $n \rightarrow \infty$ y la distribución de $(X_n - b_n)/a_n$ se aproxima por un miembro de la DVEG, entonces la distribución de X_n puede aproximarse por otro miembro de la DVEG.

La TVE se desarrolló inicialmente para una muestra aleatoria de máximos de la distribución F . Sin embargo, de acuerdo con Coles (2001), sujetos a ciertas condiciones, aún cuando la independencia temporal no sea una suposición realista, es posible suponer la independencia de eventos extremos cuando éstos se encuentran suficientemente alejados en el tiempo.

3.3 Inferencia en la DVEG.

En la práctica, la DVEG se aplica agrupando los datos en bloques de observaciones (generalmente períodos de tiempo), extrayendo el máximo muestral de cada bloque y ajustando la DVEG a la muestra de máximos resultantes, ver la Tabla 3-1.

Tabla 3-1. Muestra de máximos basada en máximos muestrales de m períodos de tiempo

Período	Datos de los períodos	Máximos
1	$Y_{11}, Y_{12}, \dots, Y_{1n_1}$	$X_1 = \max\{Y_{1j}\}$
\vdots	\vdots	\vdots
M	$Y_{m1}, Y_{m2}, \dots, Y_{mn_m}$	$X_m = \max\{Y_{mj}\}$

Por ejemplo, $Y_{j1}, Y_{j2}, \dots, Y_{jn_j}$ podrían ser las n_j mediciones diarias de cierta variable correspondientes al mes j y entonces $X_j = \max\{Y_{ij}\}$ sería el máximo del mes j . La DVEG se ajustaría a la muestra resultante de máximos extraídos X_1, \dots, X_m .

3.3.1 Estimación de parámetros de la DVEG.

Existen diversos métodos para estimar los parámetros de la DVEG, en este trabajo se utiliza el método de máxima verosimilitud, considerando que los resultados asintóticos estándar de consistencia, eficiencia y normalidad se aplican para estas distribuciones si $\xi > -0.5$ (Smith, 1985).

3.3.2 Estimadores de máxima verosimilitud.

Los estimadores de máxima verosimilitud de los parámetros de la DVEG son los valores de que maximizan la función de verosimilitud

$$L(\mu, \sigma, \xi) = \prod_{i=1}^m g(x_i; \mu, \sigma, \xi),$$

donde g es la función de densidad de la DVEG dada en la ec. (3-3). Maximizar a la verosimilitud es equivalente a maximizar a la log-verosimilitud (logaritmo de la verosimilitud). La log-verosimilitud para cuando $\xi \neq 0$ es:

$$\begin{aligned} \ell(\mu, \sigma, \xi) &= \log L(\mu, \sigma, \xi) \\ &= -m \log \sigma - (1 + 1/\xi) \sum_{i=1}^m \log \left[1 + \xi \left(\frac{z_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^m \left[1 + \xi \left(\frac{z_i - \mu}{\sigma} \right) \right]^{-1/\xi}, \end{aligned}$$

considerando que $1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) > 0$, para $i = 1, \dots, m$.

En la práctica, los estimadores de máxima verosimilitud de la DVEG son máximos locales que se encuentran por métodos numéricos de optimización no lineal.

3.3.3 Error estándar de los estimadores.

La incertidumbre en los parámetros estimados se cuantifica con sus errores estándar. De acuerdo a Coles (2001), sujetos a las limitaciones sobre ξ , la distribución aproximada de $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ es normal multivariada con media (μ, σ, ξ) y matriz de varianzas y covarianzas que se aproxima por la inversa de la *matriz de información observada*, $I_O(\mu, \sigma, \xi)$, que es igual a la matriz Hessiana de $-\ell(\mu, \sigma, \xi)$, evaluada en los estimadores de máxima verosimilitud (Smith, 2004).

Las varianzas o errores estándar de los estimadores de los parámetros de la DVEG se aproximan por la raíz cuadrada de los elementos de la diagonal principal de $I_O(\mu, \sigma, \xi)$.

3.4 Cuantiles de la DVEG.

Los cuantiles de la DVEG ajustada se usan para estimar los niveles extremos del proceso bajo estudio, éstos reciben también el nombre de *niveles de retorno* y de *valores en riesgo*. Si $\hat{\mu}$, $\hat{\sigma}$ y $\hat{\xi}$ son estimaciones de los parámetros de la DVEG, el cuantil $1 - p$ de la DVEG ajustada se obtiene al invertir la ec. (3-2) como:

$$\hat{x}_p = G^{-1}(p; \hat{\mu}, \hat{\sigma}, \hat{\xi}) = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left[1 - \{ -\log(1-p) \}^{-\hat{\xi}} \right], & \text{para } \hat{\xi} \neq 0, \\ \hat{\mu} - \hat{\sigma} \log\{ -\log(1-p) \} & , \text{ para } \hat{\xi} = 0. \end{cases} \quad (3-5)$$

El error estándar del cuantil \hat{x}_p se aproxima con el método delta:

$$\text{Var}(\hat{x}_p) \approx \nabla x_p^T I_O^{-1} \nabla x_p,$$

donde I_O es la matriz de información observada de $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$, y

$$\nabla x_p^T = [\partial x_p / \partial \mu, \partial x_p / \partial \sigma, \partial x_p / \partial \xi], \text{ evaluado en } (\hat{\mu}, \hat{\sigma}, \hat{\xi}).$$

Una vez que se estiman los errores estándar de los cuantiles estimados, se obtienen intervalos de confianza asintóticos. Así, un intervalo de confianza con un nivel de confianza del $(1-\alpha) \times 100\%$, está dado por $\hat{x}_p \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{x}_p(t))}$.

Los cuantiles \hat{x}_p permiten expresar los modelos probabilistas en unidades físicas. Éstos se interpretan como el nivel que podría ser excedido, con una probabilidad p , por el valor máximo anual de un año en particular. Por tanto, una gráfica de los cuantiles \hat{x}_p , respecto a $y = -\log(1-p)$, en escala logarítmica, denominado *gráfico de niveles de retorno*, resulta adecuada para presentar el modelo de la DVEG.

Este gráfico permite interpretar también, el comportamiento del modelo con respecto a sus parámetros. Como se muestra en la Figura 3-1, el gráfico es lineal en el caso de que $\xi = 0$. Si $\xi < 0$, el gráfico es convexo con límite asintótico $\hat{x}_p \rightarrow \mu - \sigma/\xi$, conforme $p \rightarrow 0$; en cambio si $\xi > 0$, el gráfico es cóncavo y carece de límite finito.

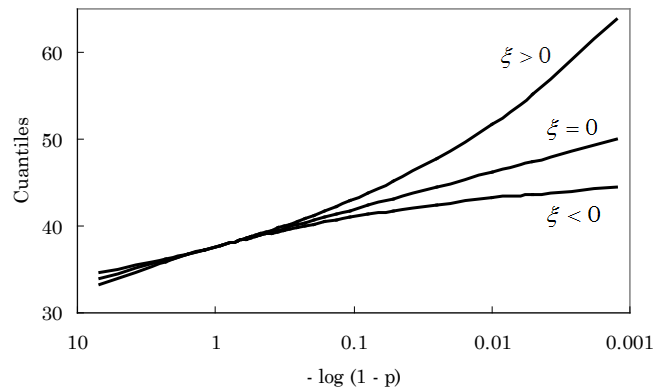


Figura 3-1. Gráfico de niveles de retorno.

3.5 Modelos que incorporan tendencia.

Cuando la muestra de máximos, X_1, \dots, X_m , contiene una tendencia, ésta se modela, de acuerdo con Coles (2001), asumiendo que el parámetro de localización, μ , de la DVEG depende del tiempo, t , y de k parámetros adicionales, β_0, \dots, β_k . Por ejemplo, ante la evidencia de una tendencia

lineal, el parámetro de localización de la DVEG podría modelarse con $\mu(t) = \beta_0 + \beta_1 t$.

A diferencia de los modelos de la sección 2.1, éste modelo es una descripción del proceso que genera los datos, y no de los datos mismos. Por lo cual se preferirán los modelos más sencillos, excepto en el caso de tener alguna evidencia teórica que respalde el uso de modelos más complejos (Coles, 2001).

Los parámetros adicionales que resultan de incluir el modelo de localización en la DVEG pueden estimarse por el método de máxima verosimilitud incluyendo a éstos en la función log-verosimilitud de la DVEG. El siguiente método, sugerido por Reiss y Thomas (1997), considera la estimación por separado de estos parámetros:

- La tendencia se estima, por mínimos cuadrados, o por algún método no paramétrico.
- Se obtiene los residuos al descontarle a los datos originales la tendencia estimada. Se ajusta la DVEG, con máxima verosimilitud, a estos residuos.
- Se deducen estimadores para, $\mu(t : \beta_0, \dots, \beta_p)$, σ y ξ .

En general, cuando se tiene una serie no estacionaria de valores extremos X_1, \dots, X_m , la estructura de la no estacionariedad puede modelarse postulando que los parámetros de la DVEG dependen del tiempo. Es decir, postulando que $X_t \sim \text{DVEG}(\mu(t), \sigma(t), \xi(t))$. Si ninguno de los parámetros de forma $\xi(t)$ es nulo, la función log-verosimilitud de la DVEG está dada por

$$\ell(\mu, \sigma, \xi) = - \sum_{t=1}^m \left\{ \log \sigma(t) + (1 + 1/\xi(t)) \log \left[1 + \xi(t) \left(\frac{x_t - \mu(t)}{\sigma(t)} \right) \right] + \left[1 + \xi(t) \left(\frac{x_t - \mu(t)}{\sigma(t)} \right) \right]^{-1/\xi(t)} \right\},$$

cuando $1 + \xi(t) \left(\frac{x_t - \mu(t)}{\sigma(t)} \right) > 0$, para $t = 1, \dots, m$.

Si además se dispone de un vector de covariables w_t para cada punto t del tiempo, podemos expresar paramétricamente al vector de parámetros de la DVEG como $\theta(t) = h(w_t^T \beta)$, donde $\theta(t) = (\mu(t), \sigma(t), \xi(t))$, h es una función, denominada “*función de liga inversa*” y β es un vector de parámetros, (Coles, 2001).

3.6 Diagnóstico del modelo.

De acuerdo con Coles (2001), para una serie de datos no estacionaria, la única forma de probar la bondad de ajuste del modelo es a través de los datos estandarizados, condicionados a los parámetros estimados. De este modo, si $X(t) \sim DVEG(\hat{\mu}(t), \hat{\sigma}(t), \hat{\xi}(t))$, la variable estandarizada $Z(t)$, definida por:

$$Z(t) = \frac{1}{\hat{\xi}(t)} \log \left\{ 1 + \hat{\xi}(t) \left(\frac{X(t) - \hat{\mu}(t)}{\hat{\sigma}(t)} \right) \right\} \quad (3-6)$$

tiene una distribución Gumbel, es decir que $P(Z(t) \leq z) = \exp(-\exp(-z))$.

Si $z_{(1)}, \dots, z_{(m)}$ denota los valores ordenados de los datos, estandarizados por la ec. (3-6), la bondad de ajuste del modelo puede verificarse por medio de un gráfica de probabilidad obtenida a partir de:

$$\left\{ (i/(m+1), \exp(-\exp(-z_{(i)}))); \quad i = 1, \dots, m \right\}. \quad (3-7)$$

Así como de un gráfico de cuantiles obtenido de:

$$\left\{ (z_{(i)}, -\log(-\log(i/(m+1))))); \quad i = 1, \dots, m \right\}. \quad (3-8)$$

CAPÍTULO 4. CASO DE ESTUDIO.

En este capítulo se aplica la Teoría de Valor Extremo al pronóstico de la demanda máxima anual de energía eléctrica, DMA, de una subestación de distribución de CFE, con base en el registro histórico de su demanda máxima mensual, DMM, durante un período de ocho años. El pronóstico se obtuvo a partir de los niveles de retorno estimados de la DVEG ajustada a los datos de la DMM, considerando las diferencias de su comportamiento en dos temporadas del año.

Al modelar la dinámica de crecimiento de la DMA en la superficie geográfica delimitada por la subestación, se consideraron dos modelos estructurales en el parámetro de localización de la DVEG, cuyos parámetros se incorporaron al modelo de la DVEG y se estimaron de manera conjunta por el método de máxima verosimilitud.

4.1 Datos de la subestación.

Las Tabla 4-1 y 4-2 muestran los valores de la demanda máxima mensual registrada en la subestación, agrupados en dos períodos de tiempo referidos como temporadas de verano e invierno: período de abril a septiembre, y el período de octubre a marzo, respectivamente.

Tabla 4-1. DMM para la temporada de invierno.

Año	Ene	Feb	Mar	Oct	Nov	Dic
1	6.25	6.25	6.25	7.93	7.94	6.66
2	6.93	6.69	8.04	15.41	9.23	9.09
3	11.70	9.80	10.11	13.40	11.72	10.86
4	11.06	10.93	11.62	17.32	14.56	11.64
5	11.64	12.04	14.28	17.64	16.88	14.16
6	13.40	13.92	15.64	20.64	23.04	23.04
7	21.48	20.88	21.36	24.36	20.04	17.52
8	14.88	16.20	20.80	24.72	20.30	18.00

Tabla 4-2. DMM para la temporada de verano.

Año	Abr	May	Jun	Jul	Ago	Sep
1	6.25	6.25	6.25	6.25	6.25	7.15
2	8.14	8.54	9.31	10.01	9.60	9.99
3	10.75	11.25	11.24	12.01	9.99	12.28
4	17.18	13.68	13.56	13.08	15.16	18.00
5	13.90	15.18	15.00	15.38	17.80	14.48
6	15.72	16.32	21.00	21.28	22.08	22.56
7	24.12	24.84	28.56	24.84	25.08	28.20
8	22.68	25.44	25.80	23.64	24.84	24.60

Estos valores se grafican en la Figura 4-1, en la cual se hace evidente el comportamiento creciente de la DMM en la subestación y su comportamiento en las temporadas consideradas.

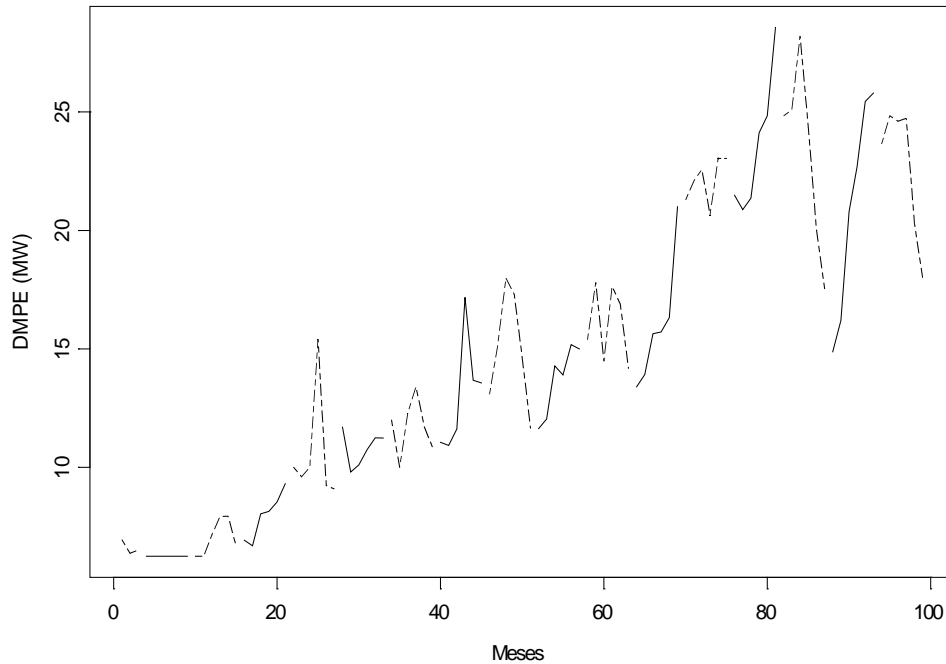


Figura 4-1. DMM en la subestación. La línea continua corresponde al período de invierno y la línea punteada corresponde al período de verano.

4.2 Modelos para el parámetro de localización.

4.2.1 Modelo lineal.

Para modelar el crecimiento de la DMM, se consideró, en primer lugar un modelo de tendencia lineal asociado al parámetro de localización de la DVEG, el cual se ajustó para cada una de las temporadas del año ($j = 1$ para invierno, $j = 2$ para verano).

$$\mu_j(t) = \beta_{0j} + \beta_{1j}t.$$

4.2.2 Modelo logístico.

El segundo modelo, diseñado con el fin de aplicar el concepto de la saturación a largo plazo de la demanda en áreas pequeñas (siendo ésta igual a la extensión del área geográfica de servicio de la subestación), utiliza el siguiente modelo de crecimiento logístico.

$$\mu_j(t) = A_j \left\{ 1 + \beta_{0j} \exp(-\beta_{1j} \times A_j \times t) \right\}^{-1} .$$

Donde A_j es un valor asintótico determinado para cada una de las temporadas del año con base en la información general disponible para este SDEE en particular.

4.3 Modelo de la DVEG con tendencia lineal en el parámetro de localización.

Considerando a X_{tjk} como una variable aleatoria que representa la DMM del mes k , de la temporada j , del año t , ($k=1,2,\dots,6$; $j=1,2$, y $t=1,2,\dots,8$), de acuerdo con la Teoría de Valor Extremo, suponemos que $X_{tjk} \sim \text{DVEG}(\mu_j(t), \sigma_j, \xi_j)$, donde $\mu_j(t) = \beta_{0j} + \beta_{1j}t$.

En este caso, la función de densidad de la DVEG en la temporada j , cuando $\xi \neq 0$, toma la forma:

$$g(x; \mu_j(t), \sigma_j, \xi_j) = \frac{1}{\sigma} \left[1 + \xi_j \left(\frac{x - \mu_j(t)}{\sigma_j} \right) \right]^{-[1+(1/\xi_j)]} \times \exp \left\{ - \left[1 + \xi_j \left(\frac{x - \mu_j(t)}{\sigma_j} \right) \right]^{-1/\xi_j} \right\}.$$

A partir de la cual se obtiene la función log-verosimilitud del modelo siguiente:

$$\ell(\beta_{01}, \beta_{11}, \sigma_1, \xi_1, \beta_{02}, \beta_{12}, \sigma_2, \xi_2) = \log \left(\prod_{j=1}^2 \prod_{t=1}^8 \prod_{k=1}^6 g(x_{jtk}; \beta_{0j}, \beta_{1j}, \sigma_j, \xi_j) \right).$$

$$\begin{aligned} \ell(\beta_{01}, \beta_{11}, \sigma_1, \xi_1, \beta_{02}, \beta_{12}, \sigma_2, \xi_2) = & \\ & - \sum_{t=1}^8 \sum_{k=1}^6 \left\{ \log \sigma_1 + \left(1 + \frac{1}{\xi_1} \right) \log \left[1 + \xi_1 \left(\frac{x_{1tk} - (\beta_{01} + \beta_{11}t)}{\sigma_1} \right) \right] \right. \\ & \left. + \left[1 + \xi_1 \left(\frac{x_{1tk} - (\beta_{01} + \beta_{11}t)}{\sigma_1} \right) \right]^{-1/\xi_1} \right\} \\ & - \sum_{t=1}^8 \sum_{k=1}^6 \left\{ \log \sigma_2 + \left(1 + \frac{1}{\xi_2} \right) \log \left[1 + \xi_2 \left(\frac{x_{2tk} - (\beta_{02} + \beta_{12}t)}{\sigma_2} \right) \right] \right. \\ & \left. + \left[1 + \xi_2 \left(\frac{x_{2tk} - (\beta_{02} + \beta_{12}t)}{\sigma_2} \right) \right]^{-1/\xi_2} \right\}, \\ \text{donde } & 1 + \xi_j \left(\frac{x_{jt}(t) - (\beta_{0j} + \beta_{1j}t)}{\sigma_j} \right) > 0; \quad j = 1, 2, \text{ y } t = 1, \dots, 8. \end{aligned}$$

4.3.1 Estimación de los parámetros de la DVEG.

La maximización numérica de la función log-verosimilitud, obtenida con el software estadístico S-Plus, produce las estimaciones de los parámetros que se muestran en la Tabla 4-3.

Tabla 4-3. Estimaciones de máxima verosimilitud de los parámetros del modelo, (errores estándares en paréntesis).

Temporada de invierno	Temporada de verano
$\hat{\beta}_{01} = 5.415 (0.827)$	$\hat{\beta}_{02} = 3.163 (0.746)$
$\hat{\beta}_{11} = 1.632 (0.218)$	$\hat{\beta}_{12} = 2.650 (0.169)$
$\hat{\sigma}_1 = 2.210 (0.346)$	$\hat{\sigma}_2 = 1.863 (0.221)$
$\hat{\xi}_1 = 0.037 (0.210)$	$\hat{\xi}_2 = -0.125 (0.111)$

La matriz aproximada de varianzas y covarianzas de los parámetros estimados se muestra en la Tabla 4-4, y la raíz cuadrada de los elementos de la diagonal principal, que corresponde al error estándar de los parámetros, (indicado entre paréntesis en la Tabla 4-3).

Tabla 4-4. Matriz aproximada de varianzas y covarianzas de los parámetros estimados de la DVEG.

Temporada de invierno	Temporada de verano
$V_1 = \begin{pmatrix} 0.682 & -0.161 & -0.113 & 0.107 \\ -0.161 & 0.047 & 0.046 & -0.037 \\ -0.113 & 0.046 & 0.119 & -0.046 \\ 0.107 & -0.037 & -0.046 & 0.044 \end{pmatrix}$	$V_2 = \begin{pmatrix} 0.556 & -0.116 & -0.047 & 0.041 \\ -0.116 & 0.029 & 0.014 & -0.012 \\ -0.047 & 0.014 & 0.049 & -0.013 \\ 0.041 & -0.012 & -0.013 & 0.012 \end{pmatrix}$

Estas matrices se obtienen observando que la función log-verosimilitud puede descomponerse en la suma de dos funciones, cada una de las cuales depende únicamente de los parámetros de una de las temporadas del año:

$$\ell(\beta_{01}, \beta_{11}, \sigma_1, \xi_1, \beta_{02}, \beta_{12}, \sigma_2, \xi_2) = \ell_1(\beta_{01}, \beta_{11}, \sigma_1, \xi_1) + \ell_2(\beta_{02}, \beta_{12}, \sigma_2, \xi_2) .$$

De este modo, la matriz de información observada del modelo puede escribirse como:

$$I_O(\beta_{01}, \beta_{11}, \sigma_1, \xi_1, \beta_{02}, \beta_{12}, \sigma_2, \xi_2) = \begin{pmatrix} I_{O_1}(\beta_{01}, \beta_{11}, \sigma_1, \xi_1) & 0 \\ 0 & I_{O_2}(\beta_{02}, \beta_{12}, \sigma_2, \xi_2) \end{pmatrix}$$

donde, ($j = 1, 2$):

$$I_{O_j}(\beta_{0j}, \beta_{1j}, \sigma_j, \xi_j) = \begin{pmatrix} -\frac{\partial^2 \ell_j}{\partial \beta_{0j}^2} & -\frac{\partial^2 \ell_j}{\partial \beta_{1j} \partial \beta_{0j}} & -\frac{\partial^2 \ell_j}{\partial \sigma_j \partial \beta_{0j}} & -\frac{\partial^2 \ell_j}{\partial \xi_j \partial \beta_{0j}} \\ -\frac{\partial^2 \ell_j}{\partial \beta_{0j} \partial \beta_{1j}} & -\frac{\partial^2 \ell_j}{\partial \beta_{1j}^2} & -\frac{\partial^2 \ell_j}{\partial \sigma_j \partial \beta_{1j}} & -\frac{\partial^2 \ell_j}{\partial \xi_j \partial \beta_{1j}} \\ -\frac{\partial^2 \ell_j}{\partial \beta_{0j} \partial \sigma_j} & -\frac{\partial^2 \ell_j}{\partial \beta_{1j} \partial \sigma_j} & -\frac{\partial^2 \ell_j}{\partial \sigma_j^2} & -\frac{\partial^2 \ell_j}{\partial \xi_j \partial \sigma_j} \\ -\frac{\partial^2 \ell_j}{\partial \beta_{0j} \partial \xi_j} & -\frac{\partial^2 \ell_j}{\partial \beta_{1j} \partial \xi_j} & -\frac{\partial^2 \ell_j}{\partial \sigma_j \partial \xi_j} & -\frac{\partial^2 \ell_j}{\partial \xi_j^2} \end{pmatrix}^{-1}$$

4.3.2 Diagnóstico del modelo.

La Figura 4-2 muestra los gráficos de probabilidad y de cuantiles utilizados para validar la bondad de ajuste del modelo.

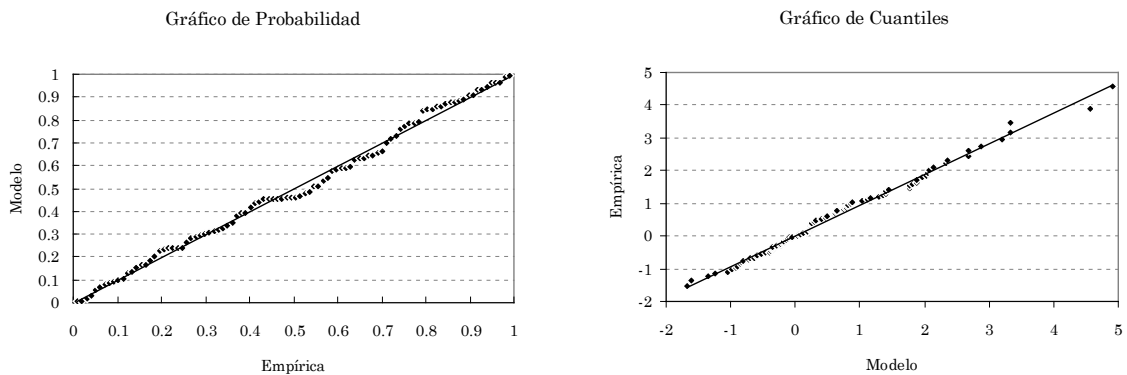


Figura 4-2. Diagnóstico del ajuste del modelo de la DVEG con tendencia lineal.

La relación razonablemente lineal observada entre el modelo ajustado y la distribución empírica de los datos muestra la validez del modelo propuesto para describir el comportamiento de la DMM en ambas temporadas del año.

4.3.3 Cuantiles de la distribución de la DMA.

El pronóstico de la DMM en cada temporada de los años futuros se obtiene a partir de la expresión del cuantil $1 - p$ de la DVEG ajustada, la cual se muestra a continuación para $t = \{9, 10, \dots\}$ y $0 < p < 1$

$$\hat{x}_{(1-p)j}(t) = \hat{\beta}_{0j} + \hat{\beta}_{1j} \cdot t - \frac{\hat{\sigma}_j}{\hat{\xi}_j} \left[1 - \{ -\log(1-p) \}^{-\hat{\xi}_j} \right], \text{ para } \hat{\xi}_j \neq 0 \text{ y } j = 1, 2.$$

Así, el cuantil $1 - p$ de la distribución de la DMM para la temporada de invierno del año t es:

$$\hat{x}_{1-p}(t) = (5.415 + 1.632t) - \frac{2.21}{0.37} \times \left\{ 1 - \left[-(\log(1-p))^{-0.37} \right] \right\};$$

y para la temporada de verano:

$$\hat{x}_{1-p}(t) = (3.163 + 2.65t) - \frac{1.863}{0.125} \times \left\{ 1 - \left[-(\log(1-p))^{0.125} \right] \right\};$$

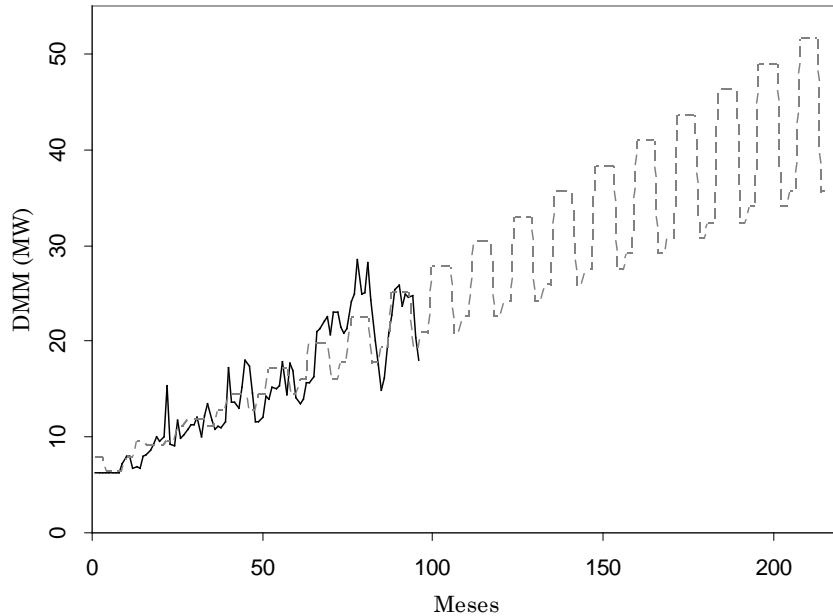


Figura 4-3. Cuantil 0.5 de la distribución estimada de la DMA para una dinámica de crecimiento lineal sin límite.

La Figura 4-3 muestra la gráfica del cuantil 0.5 de la DMM pronosticada usando la DVEG, para cada temporada de los años futuros. Se observa que el modelo captura la tendencia anual creciente, así como el diferente comportamiento de la DMM en cada una de ellas.

Dado que la DMM en verano es consistentemente mayor que la de invierno, la demanda máxima anual de la subestación, DMA, se asocia con los cuantiles estimados para la temporada de verano. Éstos se grafican en la Figura 4-4 para los años 11, 13, 15, 18, 23 y 28, y valores de p comprendidos entre 0.001 y 0.9.

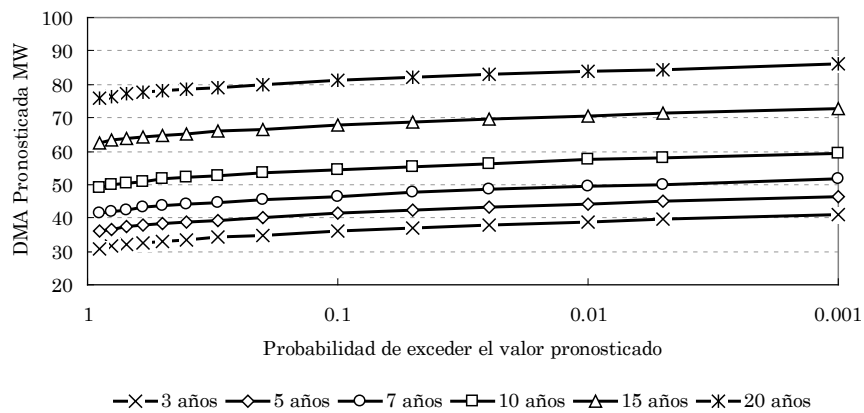


Figura 4-4. Pronóstico de la DMA, para la temporada de verano de los años 3, 5, 10, 15 y 20 futuros respecto a la probabilidad de exceder el valor pronosticado.

Los cuantiles $1 - p$ representan el nivel que sería excedido por la DMM en un año en particular, con una probabilidad p . Por tanto, si la

capacidad de la subestación es, por ejemplo, 47.5 MW (equivalentes a 50 MVA, y factor de potencia de 0.95), este modelo implica que la capacidad no sería excedida en un período de 5 años, excepto con una probabilidad menor al 0.1 %.

La probabilidad de exceder la capacidad de la subestación se incrementa evidentemente con el incremento del tiempo. Por ejemplo, de acuerdo con el modelo, existe una probabilidad tan solo del 5% de saturar la subestación en un período de 7 años, pero ésta aumenta significativamente a 30% para los 8 años futuros.

El efecto asintótico observado en la Figura 4-4 para valores de p muy pequeños, es consecuencia del valor negativo del parámetro de forma estimado para la temporada de verano, $\hat{\xi}_2$. Esto implica una distribución acotada de la DMA, y gran precisión en los cuantiles estimados, para valores de p pequeños. Sin embargo, dado que el intervalo de confianza al 95% para este parámetro, $[-0.343, 0.093]$, se extiende por arriba del cero, la evidencia a favor de una distribución acotada, no es tan fuerte.

En contraste, el gráfico de los niveles de retorno del modelo ajustado para la temporada de invierno (Figura 4-5), muestra un

comportamiento aproximadamente lineal debido al valor cercano a cero de su parámetro de forma, dando lugar a una menor precisión de los cuantiles respecto al modelo ajustado para la temporada de verano.

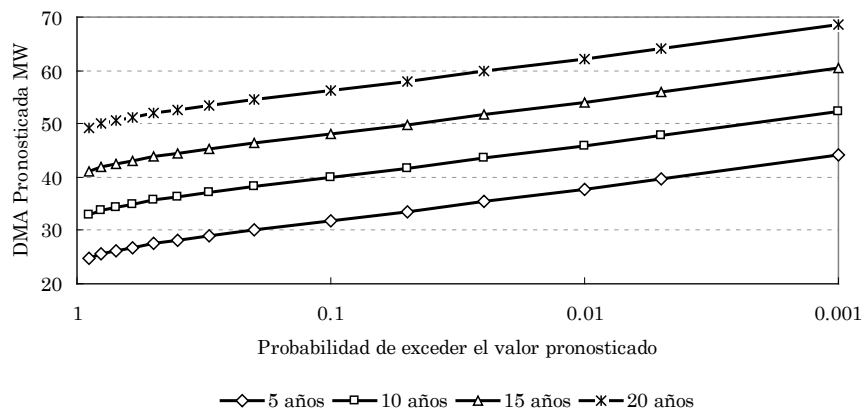


Figura 4-5. Pronóstico de la DMA, para la temporada de invierno de los años 5, 10, 15 y 20 futuros a la probabilidad de exceder el valor pronosticado.

4.3.4 Errores estándar e intervalos de confianza de los cuantiles de la DMM.

El error estándar de los cuantiles se obtiene de la raíz cuadrada de la varianza aproximada por el método delta, $Var(\hat{x}_{pj}(t)) \approx \nabla x_{pj}^T V_j \nabla x_{pj}$, donde V_j es la matriz de varianzas y covarianzas de los parámetros de la DVEG para la temporada j , y

$$\nabla x_{pj}^T = \left[\frac{\partial x_{pj}}{\partial \beta_{0j}}, \frac{\partial x_{pj}}{\partial \beta_{1j}}, \frac{\partial x_{pj}}{\partial \sigma_j}, \frac{\partial x_{pj}}{\partial \xi_j} \right]$$

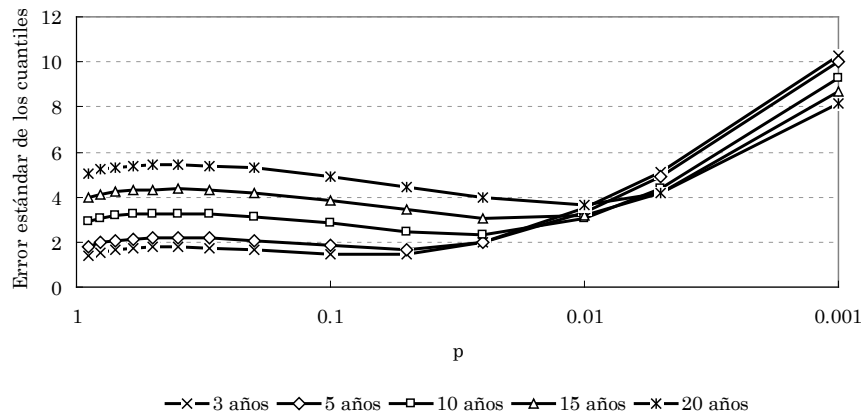
$$\nabla x_{pj}^T = \left[1, t, -\hat{\xi}_j^{-1} \left\{ 1 - (-\log(1-p))^{-\hat{\xi}_j} \right\}, \right. \\ \left. \hat{\sigma}_j \hat{\xi}_j^{-2} \left\{ 1 - (-\log(1-p))^{-\hat{\xi}_j} \right\} - \hat{\sigma}_j \hat{\xi}_j^{-1} (-\log(1-p))^{-\hat{\xi}_j} \log(-\log(1-p)) \right].$$

La Tabla 4-5 muestra el cálculo de este gradiente para valores de $p = 0.5, 0.1$ y 0.01 .

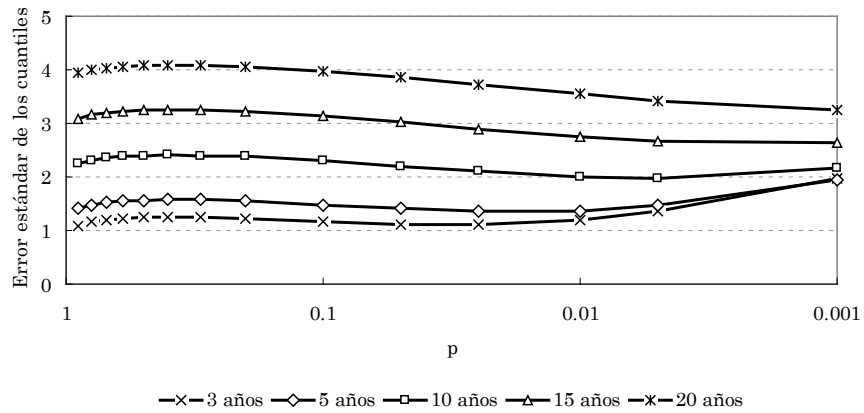
Tabla 4-5. Gradiente asociado al cuantil $(1-p)$ de la DVEG.

Temporada de invierno	Temporada de verano
$\nabla x_{0.5}^T = [1, t, 0.369 \quad 0.1498]$	$\nabla x_{0.5}^T = [1, t, 0.358 \quad 0.1214]$
$\nabla x_{0.9}^T = [1, t, 2.347 \quad 5.916]$	$\nabla x_{0.9}^T = [1, t, 1.962 \quad 3.919]$
$\nabla x_{0.99}^T = [1, t, 5.015 \quad 26.214]$	$\nabla x_{0.99}^T = [1, t, 3.498 \quad 13.562]$

Los errores estándar de los cuantiles estimados se grafican en la Figura 4-6, para $0.001 \leq p \leq 0.9$. Se observa poca variabilidad de estos cuantiles para la temporada de verano, en cambio, es evidente el incremento en la incertidumbre en los cuantiles del período de invierno para $p < 0.005$.



(a) Temporada de Invierno.



(b) Temporada de verano.

Figura 4-6. Gráfica del error estándar de los cuantiles, con un horizonte de pronóstico a 20 años, usando el modelo de tendencia lineal.

Tabla 4-6 y 4-7 muestran el error estándar y los intervalos de confianza al 95%, calculados para los cuantiles estimados, \hat{x}_p , con un horizonte de pronóstico de 20 años, y valores de $p = 0.5, 0.1, 0.01$.

*Tabla 4-6. DMA pronosticada en la subestación,
(temporada de invierno).*

T	$\hat{x}_{0.5}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.9}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.99}$	Err. Est.	I.C. 95% Conf.
9	20.92	(1.34)	(18.3, 23.5)	25.29	(1.15)	(23.0, 27.5)	31.19	(3.74)	(23.9, 38.5)
10	22.55	(1.55)	(19.5, 25.6)	26.92	(1.29)	(24.4, 29.5)	32.82	(3.62)	(25.7, 39.9)
11	24.18	(1.76)	(20.7, 27.6)	28.55	(1.46)	(25.7, 31.4)	34.45	(3.51)	(27.6, 41.3)
12	25.81	(1.97)	(22.0, 29.7)	30.19	(1.64)	(27.0, 33.4)	36.08	(3.41)	(29.4, 42.8)
13	27.45	(2.18)	(23.2, 31.7)	31.82	(1.82)	(28.2, 35.4)	37.71	(3.32)	(31.2, 44.2)
14	29.08	(2.40)	(24.4, 33.8)	33.45	(2.02)	(29.5, 37.4)	39.35	(3.24)	(33.0, 45.7)
15	30.71	(2.61)	(25.6, 35.8)	35.08	(2.21)	(30.7, 39.4)	40.98	(3.18)	(34.7, 47.2)
16	32.34	(2.83)	(26.8, 37.9)	36.71	(2.41)	(32.0, 41.4)	42.61	(3.13)	(36.5, 48.7)
17	33.97	(3.04)	(28.0, 39.9)	38.35	(2.62)	(33.2, 43.5)	44.24	(3.09)	(38.2, 50.3)
18	35.61	(3.26)	(29.2, 42.0)	39.98	(2.82)	(34.4, 45.5)	45.87	(3.07)	(39.9, 51.9)
19	37.24	(3.47)	(30.4, 44.0)	41.61	(3.03)	(35.7, 47.6)	47.51	(3.07)	(41.5, 53.5)
20	38.87	(3.69)	(31.6, 46.1)	43.24	(3.24)	(36.9, 49.6)	49.14	(3.08)	(43.1, 55.2)
21	40.50	(3.91)	(32.8, 48.2)	44.87	(3.45)	(38.1, 51.6)	50.77	(3.11)	(44.7, 56.9)
22	42.13	(4.12)	(34.1, 50.2)	46.51	(3.66)	(39.3, 53.7)	52.40	(3.15)	(46.2, 58.6)
23	43.77	(4.34)	(35.3, 52.3)	48.14	(3.87)	(40.5, 55.7)	54.03	(3.20)	(47.8, 60.3)
24	45.40	(4.56)	(36.5, 54.3)	49.77	(4.08)	(41.8, 57.8)	55.67	(3.27)	(49.3, 62.1)
25	47.03	(4.77)	(37.7, 56.4)	51.40	(4.30)	(43.0, 59.8)	57.30	(3.35)	(50.7, 63.9)
26	48.66	(4.99)	(38.9, 58.4)	53.03	(4.51)	(44.2, 61.9)	58.93	(3.44)	(52.2, 65.7)
27	50.29	(5.21)	(40.1, 60.5)	54.67	(4.72)	(45.4, 63.9)	60.56	(3.55)	(53.6, 67.5)
28	51.93	(5.43)	(41.3, 62.6)	56.30	(4.94)	(46.6, 66.0)	62.19	(3.66)	(55.0, 69.4)

*Tabla 4-7. DMA pronosticada en la subestación,
(temporada de verano).*

T	$\hat{x}_{0.5}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.9}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.99}$	Err. Est.	I.C. 95% Conf.
9	27.68	(0.91)	(25.9, 29.5)	30.67	(0.87)	(29.0, 32.4)	33.53	(1.09)	(31.4, 35.7)
10	30.33	(1.07)	(28.2, 32.4)	33.32	(1.02)	(31.3, 35.3)	36.18	(1.13)	(34.0, 38.4)
11	32.98	(1.24)	(30.6, 35.4)	35.97	(1.17)	(33.7, 38.3)	38.83	(1.19)	(36.5, 41.2)
12	35.63	(1.40)	(32.9, 38.4)	38.62	(1.32)	(36.0, 41.2)	41.48	(1.27)	(39.0, 44.0)
13	38.28	(1.57)	(35.2, 41.4)	41.27	(1.48)	(38.4, 44.2)	44.13	(1.37)	(41.5, 46.8)

Tabla 4-7.Cont...

T	$\hat{x}_{0.5}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.9}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.99}$	Err. Est.	I.C. 95% Conf.
14	40.93	(1.73)	(37.5, 44.3)	43.92	(1.64)	(40.7, 47.1)	46.78	(1.48)	(43.9, 49.7)
15	43.58	(1.90)	(39.9, 47.3)	46.57	(1.81)	(43.0, 50.1)	49.43	(1.60)	(46.3, 52.6)
16	46.23	(2.07)	(42.2, 50.3)	49.22	(1.97)	(45.4, 53.1)	52.08	(1.72)	(48.7, 55.5)
17	48.88	(2.23)	(44.5, 53.3)	51.87	(2.13)	(47.7, 56.0)	54.73	(1.86)	(51.1, 58.4)
18	51.53	(2.40)	(46.8, 56.2)	54.52	(2.30)	(50.0, 59.0)	57.38	(2.00)	(53.5, 61.3)
19	54.18	(2.57)	(49.1, 59.2)	57.17	(2.46)	(52.3, 62.0)	60.03	(2.14)	(55.8, 64.2)
20	56.83	(2.74)	(51.5, 62.2)	59.82	(2.63)	(54.7, 65.0)	62.68	(2.29)	(58.2, 67.2)
21	59.48	(2.90)	(53.8, 65.2)	62.47	(2.79)	(57.0, 67.9)	65.33	(2.44)	(60.5, 70.1)
22	62.13	(3.07)	(56.1, 68.2)	65.12	(2.96)	(59.3, 70.9)	67.98	(2.59)	(62.9, 73.1)
23	64.78	(3.24)	(58.4, 71.1)	67.77	(3.13)	(61.6, 73.9)	70.63	(2.75)	(65.2, 76.0)
24	67.43	(3.41)	(60.8, 74.1)	70.42	(3.30)	(64.0, 76.9)	73.28	(2.91)	(67.6, 79.0)
25	70.08	(3.58)	(63.1, 77.1)	73.07	(3.46)	(66.3, 79.9)	75.93	(3.06)	(69.9, 81.9)
26	72.73	(3.74)	(65.4, 80.1)	75.72	(3.63)	(68.6, 82.8)	78.58	(3.22)	(72.3, 84.9)
27	75.38	(3.91)	(67.7, 83.1)	78.37	(3.80)	(70.9, 85.8)	81.23	(3.38)	(74.6, 87.9)
28	78.03	(4.08)	(70.0, 86.0)	81.02	(3.96)	(73.2, 88.8)	83.88	(3.54)	(76.9, 90.8)

4.3.5 Pronóstico de la DMA de la subestación.

Con base en estos resultados pueden establecerse algunos criterios para dimensionar la capacidad requerida en el área geográfica de atendida por este SDEE. En la Figura 4-7 se muestran diferentes niveles de predicción para la demanda, correspondientes a diferentes valores de probabilidad de exceder estos niveles.

De acuerdo con la figura, pueden establecerse límites de diseño tan robustos como sea necesario con el fin de determinar la capacidad requerida por el SDDE. Si se elige como límite de diseño el determinado por los cuantiles 0.999, entonces se tendría un riesgo muy pequeño, del orden del 0.1 % de exceder el valor pronosticado.

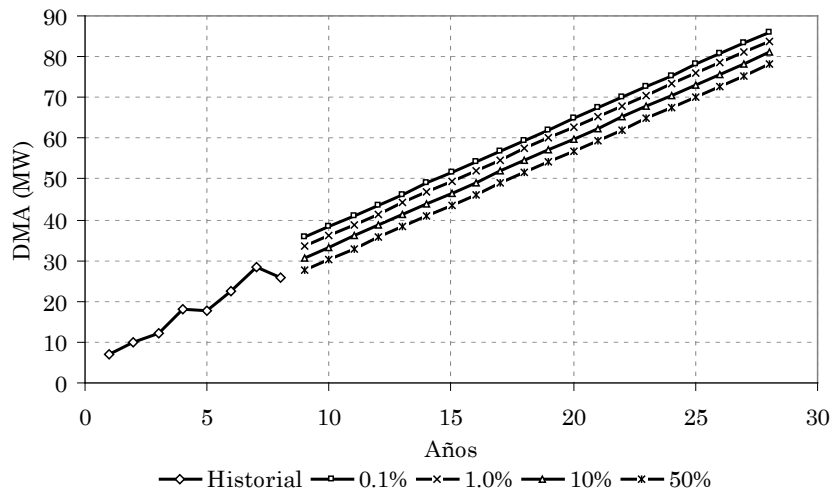


Figura 4-7. Pronóstico a 20 años de la DMA de la subestación, con una probabilidad de 0.1%, 1%, 10% y 50% de exceder el valor pronosticado.

Con base en este pronóstico y considerando una capacidad instalada en la subestación de 45 MW, el planificador de este SDEE podría establecer los siguientes criterios de diseño:

- Las obras de expansión necesarias para solucionar los problemas de saturación en esta subestación deberían entrar

en operación a lo más en el año 13 (año 5 del horizonte de pronóstico).

- Para el año 23 se requiere una capacidad instalada mayor o igual a 73 MW (≈ 80 MVA).
- Se requiere una capacidad instalada en el sistema mayor o igual a 86 MW (≈ 100 MVA) al final de los 20 años.

4.4 Modelo de la DVEG con tendencia logística en el parámetro de localización.

Para ilustrar la forma de incluir los modelos de áreas pequeñas en los modelos de valores extremos, se pronosticó la DMA del período de verano para un horizonte de 20 años, considerando el modelo de crecimiento logístico simple. Siguiendo el procedimiento descrito en la sección 4.3, se maximizó la función log-verosimilitud de la DVEG reemplazando la relación de dependencia del parámetro de localización por el siguiente modelo logístico:

$$\mu(t) = A \{1 + \beta_0 \exp(-\beta_1 \cdot A \cdot t)\}^{-1}.$$

La función log-verosimilitud del modelo, en este caso es la siguiente:

$$\ell(\beta_0, \beta_1, \sigma, \xi) =$$

$$-\sum_{t=1}^8 \sum_{k=1}^6 \left\{ \log \sigma + \left(1 + \frac{1}{\xi}\right) \log \left[1 + \xi \left(\frac{x_{tk} - \left(A \{1 + \beta_0 \exp(-\beta_1 \cdot A \cdot t)\}^{-1} \right)}{\sigma} \right) \right] \right. \\ \left. + \left[1 + \xi \left(\frac{x_{tk} - \left(A \{1 + \beta_0 \exp(-\beta_1 \cdot A \cdot t)\}^{-1} \right)}{\sigma} \right) \right]^{-1/\xi} \right\},$$

$$\text{donde } 1 + \xi \left(\frac{x_t(t) - \left(A \{1 + \beta_0 \exp(-\beta_1 \cdot A \cdot t)\}^{-1} \right)}{\sigma} \right) > 0; \quad t = 1, \dots, 8.$$

4.4.1 Parámetros estimados de la DVEG.

Antes de maximizar la función log-verosimilitud fue necesario determinar un valor asintótico razonable para la demanda del área delimitada por la subestación: considerando un radio de influencia de la subestación de 2,5 km, y una densidad de carga promedio en la región de 0.05 MW/ha, el desarrollo de esta “pequeña área”, cuya extensión aproximada es de 2 000 has, se saturaría alrededor de 100 MW.

En la Tabla 4-8 se muestran los parámetros ajustados para el modelo de valores extremos.

Tabla 4-8. Estimaciones de máxima verosimilitud de los parámetros del modelo, (errores estándares en paréntesis).

Parámetros de localización	Parámetros de escala y forma
$\hat{\beta}_0 = 20.233 (3.21)$	$\hat{\sigma} = 1.9999 (0.468)$
$\hat{\beta}_1 = 0.00214 (0.00024)$	$\hat{\xi} = 0.09989 (0.413)$

Y a partir de la matriz de información observada, evaluada en $(\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}, \hat{\xi})$:

$$I_O(\beta_0, \beta_1, \sigma, \xi) = \begin{pmatrix} -\frac{\partial^2 l}{\partial \beta_0^2} & -\frac{\partial^2 l}{\partial \beta_1 \partial \beta_0} & -\frac{\partial^2 l}{\partial \sigma \partial \beta_0} & -\frac{\partial^2 l}{\partial \xi \partial \beta_0} \\ -\frac{\partial^2 l}{\partial \beta_0 \partial \beta_1} & -\frac{\partial^2 l}{\partial \beta_1^2} & -\frac{\partial^2 l}{\partial \sigma \partial \beta_1} & -\frac{\partial^2 l}{\partial \xi \partial \beta_1} \\ -\frac{\partial^2 l}{\partial \beta_0 \partial \sigma} & -\frac{\partial^2 l}{\partial \beta_1 \partial \sigma} & -\frac{\partial^2 l}{\partial \sigma^2} & -\frac{\partial^2 l}{\partial \xi \partial \sigma} \\ -\frac{\partial^2 l}{\partial \beta_0 \partial \xi} & -\frac{\partial^2 l}{\partial \beta_1 \partial \xi} & -\frac{\partial^2 l}{\partial \sigma \partial \xi} & -\frac{\partial^2 l}{\partial \xi^2} \end{pmatrix}^{-1}$$

Se obtiene la siguiente matriz de varianzas y covarianzas:

$$V = \begin{pmatrix} 10.307190 & 0.000768 & 1.023964 & -1.097416 \\ 0.000768 & 0.00000006 & 0.000084 & -0.000087 \\ 1.023964 & 0.000084 & 0.218636 & -0.138193 \\ -1.097416 & -0.000087 & -0.138193 & 0.170557 \end{pmatrix}$$

4.4.2 Diagnóstico del modelo.

La Figura 4-8 muestra las gráficas de probabilidad y de cuantiles, utilizadas para validar la bondad de ajuste de este modelo. Ambos gráficos muestra un comportamiento razonablemente lineal, con excepción del gráfico de cuantiles para valores mayores de z , que corresponden a la temporada de invierno.

Se concluye, por tanto, que el modelo propuesto describe adecuadamente el comportamiento de la demanda máxima mensual en la temporada de verano.

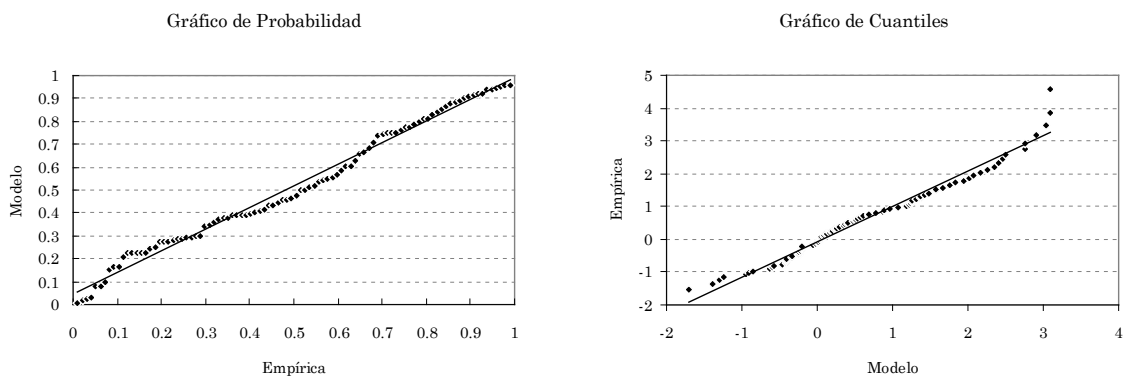


Figura 4-8. Diagnóstico del ajuste del modelo de la DVEG con tendencia logística.

4.4.3 Cuantiles de la distribución de la DMA.

La siguiente ecuación define el cuantil $(1 - p)$ de la distribución de la DMA para los años futuros:

$$\hat{x}_p(t) = A \{1 + \hat{\beta}_0 \exp(-\hat{\beta}_1 \cdot A \cdot t)\}^{-1} - \frac{\hat{\sigma}}{\hat{\xi}} \left[1 - \{ -\log(1 - p) \}^{-\hat{\xi}} \right], \quad (4-1)$$

Sustituyendo los parámetros estimados de la Tabla 4-8 se tiene que:

$$\hat{x}_p(t) = 100 \{1 + 20.23 \exp(-0.214 \cdot t)\}^{-1} - 20.02 \left[1 - \{ -\log(1 - p) \}^{-0.1} \right],$$

La Figura 4-9 muestra el pronóstico de la DMA con respecto a la probabilidad de exceder el valor pronosticado.

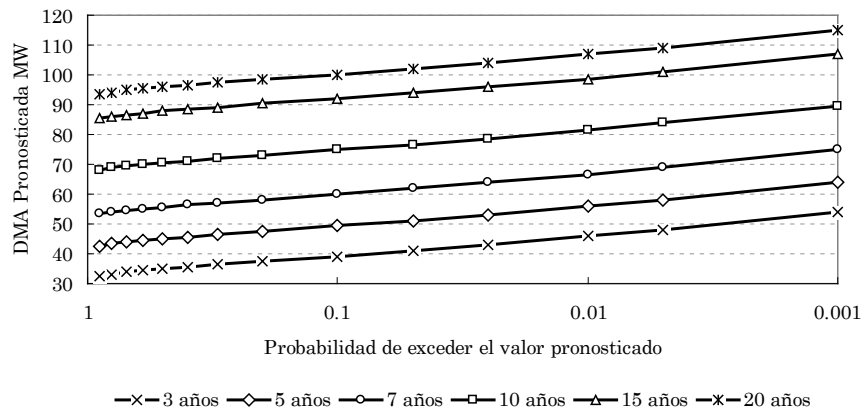


Figura 4-9. Pronóstico de la DMA, para los años 3, 5, 7, 10, 15 y 20 futuros, en función de la probabilidad de exceder el valor pronosticado.

Se indican los niveles de retorno de la distribución de la DMA pronosticada para los años 11, 13, 15, 18, 23 y 28, obtenidos para un rango de valores de p comprendido entre 0.001 y 0.5. Esta gráfica permite considerar el nivel de riesgo asociado con el nivel de retorno elegido para el pronóstico, de este modo, de acuerdo con el modelo, si se elige un nivel de riesgo del 1%, se espera que la demanda del año 5 sea menor o igual a 55 MW.

4.4.4 Cálculo de los errores estándar.

La siguiente expresión muestra el gradiente de la ec. (4-1),

$$\nabla x_p = \begin{bmatrix} -A \exp(-\hat{\beta}_1 A t) \{1 + \hat{\beta}_0 \exp(-\hat{\beta}_1 A t)\}^{-2} \\ A^2 \hat{\beta}_0 t \exp(-\hat{\beta}_1 A t) \{1 + \hat{\beta}_0 \exp(-\hat{\beta}_1 A t)\}^{-2} \\ -\hat{\xi}^{-1} \left\{ 1 - (-\log(1-p))^{-\hat{\xi}} \right\} \\ \left[\hat{\sigma}_{\xi}^{\hat{\xi}-2} \left\{ 1 - (-\log(1-p))^{-\hat{\xi}} \right\} - \hat{\sigma}_{\xi}^{\hat{\xi}-1} (-\log(1-p))^{-\hat{\xi}} \log(-\log(1-p)) \right] \end{bmatrix}$$

A partir del cual se calculan los errores estándar de los cuantiles, aplicando el método delta $Var(\hat{x}_p(t)) \approx \nabla x_p^T V \nabla x_p$. La Figura 4-10 muestra la gran incertidumbre que se tiene para los cuantiles estimados

con $p < 0.025$. Y la Tabla 4-9 los errores estándar y los intervalos de confianza calculados de los cuantiles para valores de $p = 0.5, 0.1$ y 0.01 .

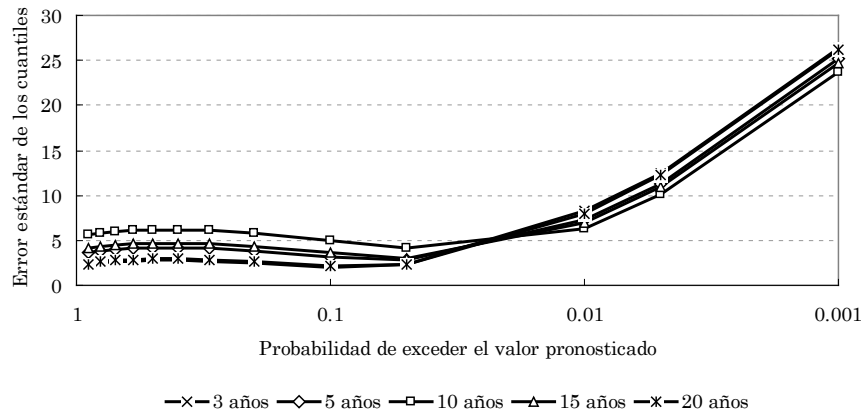


Figura 4-10. Gráfica del error estándar de los niveles de retorno para los años de pronóstico 3, 5, 10, 15 y 20, usando el modelo de tendencia logístico.

Tabla 4-9. DMA pronosticada en la subestación, (tendencia logística).

T	$\hat{x}_{0.5}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.9}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.99}$	Err. Est.	I.C. 95% Conf.
9	26.02	(1.48)	(23.1 28.9)	30.32	(1.47)	(27.4 33.2)	36.95	(9.19)	(18.9 55.0)
10	30.26	(2.09)	(26.2 34.4)	34.56	(1.63)	(31.4 37.8)	41.19	(8.71)	(24.1 58.3)
11	34.89	(2.77)	(29.5 40.3)	39.19	(2.03)	(35.2 43.2)	45.82	(8.21)	(29.7 61.9)
12	39.84	(3.49)	(33.0 46.7)	44.14	(2.59)	(39.1 49.2)	50.77	(7.72)	(35.6 65.9)
13	45.03	(4.20)	(36.8 53.3)	49.33	(3.21)	(43.0 55.6)	55.96	(7.28)	(41.7 70.2)
14	50.35	(4.86)	(40.8 59.9)	54.65	(3.81)	(47.2 62.1)	61.28	(6.91)	(47.7 74.8)
15	55.67	(5.41)	(45.1 66.3)	59.97	(4.32)	(51.5 68.4)	66.61	(6.63)	(53.6 79.6)
16	60.89	(5.82)	(49.5 72.3)	65.19	(4.71)	(56.0 74.4)	71.82	(6.45)	(59.2 84.5)
17	65.88	(6.07)	(54.0 77.8)	70.18	(4.95)	(60.5 79.9)	76.82	(6.35)	(64.4 89.3)
18	70.57	(6.15)	(58.5 82.6)	74.87	(5.02)	(65.0 84.7)	81.50	(6.31)	(69.1 93.9)
19	74.87	(6.07)	(63.0 86.8)	79.17	(4.95)	(69.5 88.9)	85.80	(6.33)	(73.4 98.2)
20	78.75	(5.86)	(67.3 90.2)	83.05	(4.74)	(73.8 92.3)	89.69	(6.41)	(77.1 102.3)

Tabla 4-9. Cont...

T	$\hat{x}_{0.5}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.9}$	Err. Est.	I.C. 95% Conf.	$\hat{x}_{0.99}$	Err. Est.	I.C. 95% Conf.
21	82.20	(5.54)	(71.3 93.1)	86.50	(4.43)	(77.8 95.2)	93.13	(6.54)	(80.3 106.0)
22	85.21	(5.14)	(75.1 95.3)	89.51	(4.05)	(81.6 97.5)	96.15	(6.72)	(83.0 109.3)
23	87.82	(4.70)	(78.6 97.0)	92.12	(3.63)	(85.0 99.2)	98.75	(6.94)	(85.1 112.4)
24	90.04	(4.24)	(81.7 98.3)	94.34	(3.21)	(88.1 100.6)	100.97	(7.20)	(86.9 115.1)
25	91.92	(3.78)	(84.5 99.3)	96.22	(2.79)	(90.7 101.7)	102.85	(7.47)	(88.2 117.5)
26	93.49	(3.33)	(87.0 100.0)	97.79	(2.41)	(93.1 102.5)	104.42	(7.75)	(89.2 119.6)
27	94.80	(2.92)	(89.1 100.5)	99.10	(2.07)	(95.0 103.2)	105.74	(8.02)	(90.0 121.5)
28	95.89	(2.54)	(90.9 100.9)	100.19	(1.79)	(96.7 103.7)	106.82	(8.29)	(90.6 123.1)

4.4.5 Pronóstico de la DMA.

El pronóstico de la DMA se muestra en la Figura 4-11, para un horizonte de 20 años, con base en los cuantiles 0.975 de la DVEG ajustada.

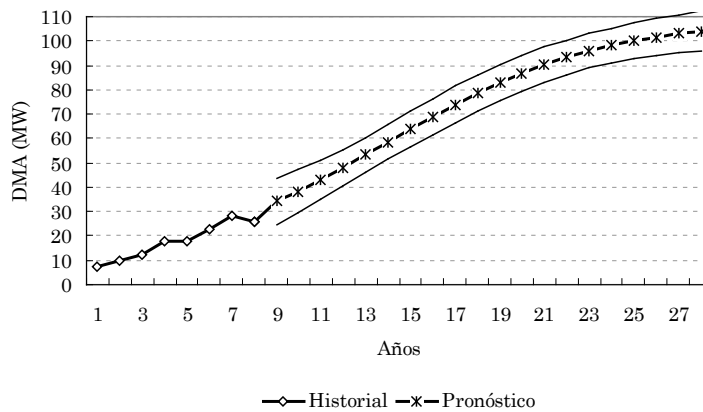


Figura 4-11. Pronóstico a 20 años de la DMA de la subestación, considerando una probabilidad del 2.5% de exceder el valor pronosticado.

Es decir, de acuerdo a la figura, se espera que la DMA sea menor o igual al valor pronosticado con una probabilidad de 97.5%. Se utiliza este grado de confianza debido a la pérdida de precisión del modelo para valores de p menores a 0.025.

Con base en este pronóstico se tendrían los siguientes criterios de planificación:

- Las obras de expansión deberán entrar en operación a lo más en el año 11 (año 3 del horizonte de pronóstico).
- Para el año 17 se requiere una capacidad instalada mayor o igual a 74 MW (≈ 80 MVA).
- Dada la saturación del área de servicio en el año 28 (año 20 del pronóstico), se requerirá una capacidad instalada mayor o igual a 104 MW (≈ 120 MVA), al final de los 20 años.

CAPÍTULO 5. CONCLUSIONES Y RECOMENDACIONES.

5.1 Conclusiones.

En este trabajo se describe un método para pronosticar la demanda máxima anual, DMA, en sistemas de distribución de energía eléctrica, con horizontes a largo plazo, mostrando su aplicación con los datos de demanda máxima mensual de una subestación de distribución de la CFE.

El método propuesto, fundamentado en la Teoría de Valor Extremo, descrita en el Capítulo 3, permite incluir, de manera flexible, diferentes dinámicas de crecimiento relacionadas con el desarrollo de áreas pequeñas, como las que se describen en el Capítulo 2, mediante la elección de un modelo estructural asociado al parámetro de localización de la distribución de valor extremo generalizada.

Las pruebas de bondad de ajuste y la variabilidad de los parámetros estimados para cada uno de los modelos analizados en el Capítulo 4, muestran que los datos proporcionados por CFE son suficientes para obtener un modelo adecuado para describir el comportamiento de la DMA de la subestación.

Los cuantiles de la distribución de valor extremo generalizada de los modelos ajustados se utilizaron para definir niveles de riesgo asociados con la elección de la capacidad de diseño del sistema eléctrico. Los errores estándar obtenidos para estos cuantiles muestran que utilizando toda la información disponible en CFE, es posible reducir la incertidumbre del pronósticos, a diferencia de lo que ocurre con los métodos de extrapolación que solo utilizan un valor para cada uno de los años que componen el historial de la demanda máxima.

La elección de una curva de crecimiento para modelar la tendencia es una de las decisiones del analista que dependerá generalmente del enfoque del pronóstico. El modelo de crecimiento lineal implica el crecimiento ilimitado de la demanda en el área atendida por la subestación, lo cual puede ocurrir si los límites del área de estudio se extienden continuamente para absorber mayor demanda debida al

desarrollo demográfico. En cambio, si la extensión del área de estudio se mantiene constante, el uso del modelo logístico describirá el cambio en la tasa de crecimiento conforme el desarrollo tiende a un valor de saturación.

Independientemente del modelo estructural utilizado para describir la dinámica de crecimiento, los modelos de valores extremos proporcionan una medida de riesgo con el cual las realizaciones futuras de la DMA excederán el valor pronosticado.

Todo proceso de planeación involucra la toma de decisiones bajo situaciones de incertidumbre. Los efectos de esta incertidumbre se hacen evidentes cuando las decisiones han sido tomadas y los planes instrumentados; lo que implica que el pronóstico para la planeación de un SDEE puede llevar a decisiones equivocadas. Una forma de reducir este riesgo es mediante el uso de técnicas adecuadas para el pronóstico. En este sentido, los modelos de valores extremos deben ser considerados como otra opción para el pronóstico de la DMA.

5.2 Ventajas y desventajas.

El método propuesto considera objetivamente diferentes escenarios de pronóstico para diseñar y planificar un SDEE, proporcionando medidas

de riesgo que le permiten al planificador elegir entre distintos grados de robustez para el diseño del sistema, y modela de manera explícita la no-estacionariedad de la DMM permitiendo elegir diferentes modelos de crecimiento asociados con el parámetro de localización de la DVEG, que distinguen su comportamiento en diferentes temporadas del año.

Sin embargo, requiere de datos de períodos históricos largos y de técnicas especializadas de modelación estadística que hasta el momento son difíciles de automatizar.

5.3 Contribución.

Este trabajo contribuye a mejorar los métodos de extrapolación utilizados actualmente por CFE, al incorporar una medida de riesgo asociada a los pronósticos de la DMA, la cual resulta de gran importancia para los planificadores del sistema eléctrico al evaluar el impacto de tales decisiones.

Así mismo, al considerar que los datos de la DMM se distribuyen de acuerdo con la DVEG, la cual abarca los tres tipos de distribuciones de valor extremo: Gumbel, Frechet y Weibul, se contribuye a extender la investigación de Loredó (2001) y Belzer (1993), quienes asumieron que los

datos de la demanda máxima se distribuían de acuerdo con la distribución de valor extremo Tipo I (Gumbel) y sugirieron, como una forma de mejorar sus resultados, probar la posibilidad de que los datos se distribuyeran de acuerdo con las distribuciones de valor extremo Tipo II y Tipo III.

5.4 Investigación actual y futura.

Una de las características importantes del modelo propuesto es el supuesto de independencia entre las observaciones de la DMM. Por esta razón, debe investigarse la aplicación de modelos no-paramétricos y semi-paramétricos que se adapten con flexibilidad al comportamiento de los datos, véase por ejemplo Hall y Tajvidi (2000).

Los modelos estudiados se aplican al pronóstico de la DMA en SDEE urbanos cerrados y subestaciones rurales, en los cuales la interacción con otros sistemas se considera despreciable. En muchos casos se requiere analizar historiales de una o más subestaciones adyacentes en las cuales existe evidencia de que una de ellas transfirió a las otras, parte de su demanda en algún momento de la serie de tiempo.

Modelar y pronosticar la variable *estructural* para estos SDEE requiere modelos para estudiar sistemas que evolucionan tanto en el

tiempo como en el espacio, los llamados modelos *espacio-temporales*. Los modelos espacio-temporales para extremos es un área reciente donde actualmente se realiza investigación. Una línea de investigación posible es considerar el método de *acoplamiento de carga transferida* o LTC, por sus siglas en inglés, propuesto por Willis, Powell y Wall (1984) con el fin de proponer una alternativa al modelo estructural de los parámetros de localización de las subestaciones involucradas.

Referencias.

- Alfares, H. K. & Nazeeruddin, M. (2001). Electric load forecasting: Literature survey and classification of methods. *International Journal of Systems Science*, Vol. 33, No.1, 23-34.
- Belzer, D. B. & Kellogg, M. A. (1993). Incorporating sources of uncertainty in forecasting peak power loads: A Monte Carlo analysis using Extreme Value Distribution. *IEEE Transactions on Power Systems*, Vol 8, No. 2, 730-737.
- Bossert, R. W. (1977). The logistic growth curve reviewed, programmed, and applied to electric utility forecasting. *Technological Forecasting and Social Change*, No 10, 357-368.
- Brown, R. E., Hanson, A. P. & Hagan, D. L. (1999). Long range spatial load forecasting using non-uniform areas. *IEEE Transmission and Distribution Conference*, 11-16 April, Vol. 1, 369-373.
- Castillo, E. (1988). *Extreme Value Theory in Engineering*. New York: Academic Press, Inc.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. New York: Springer.
- Diebold, F. X. (1999). *Elementos de Pronósticos*. México, D.F.: Thomson.

- Elias, C. N. & Hatziargyriou, N. D. (2009). An annual midterm energy forecasting model using fuzzy logic. *IEEE Transactions on Power Systems*, Vol. 24, No. 1, 469-478.
- Embrechts, P., Klupelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events*. Berlin: Springer.
- Engel, M. V & Palmer, R. E. (1992). Distribution planning. En Engel, M. V., Green, E. R. & Willis, H. L. (Eds.). *IEEE Tutorial on Distribution Planning, Course Text EHO 361-6-PWR*.
- Fisher, R. A. & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest number of sample. *Proc. Cambridge Philos. Soc.*, Vol. 24, Pt 2, 180-190.
- Fu, C. W. & Nguyen, T. T. (2003). Models for long-term energy forecasting. *IEEE Power Engineering Society General Meeting*, Vol. 4 –(13-17, July 2003).
- Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. New York: John Wiley and Sons.
- Gellings, C. W. (Ed.) (1996). *Demand Forecasting in the Electric Utility Industry* (2^a Ed.), Oklahoma: PennWell.
- Genethliou, D. (2005). *Statistical Approaches to Electric Load Forecasting*. Tesis de Doctorado en Matemáticas Aplicadas y Estadística, Stony Brook University.
- Gnedenko, B. V. (1943). Sur la distribution limitè du terme d'une serie aleatoire. *Ann Math.* 44, 423-453.
- Gönen, T. (1986). *Electric Power Distribution Systems Engineering*. Singapore: McGraw-Hill.
- Gordon, (1978). *System Simulation* (2^a ed). New Jersey: Prentice Hall.
- Grossman, S. I., (1996). *Álgebra Lineal*, (5^a Ed.). México, D.F.: Mc Graw Hill.

- Gumbel, E. J. (1958). *Statistics of Extremes*. New York: Dover Publications, Inc. (Originalmente publicado por: New York: Columbia University Press).
- Haan, L. & Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer.
- Hall, P. & Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, Vol. 15, No. 2, 153-167.
- Henault, P. H., Eastvedt, R. B., Peschon, J. & Hajdu, L. P. (1970). Power system long-term planning in the presence of uncertainty. *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, No. 1, 156-164.
- Huck, G. E., Mahmoud, A. A., Comerford, R. B., Adams, J. & Dawson, E. (1980). Load forecast bibliography: Phase I. *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-99, No. 1, 53-58.
- Hyndman, R. J. & Fan, S. (2008). *Density Forecasting for Long-term Peak*. Disponible en: www.buseco.monash.edu.au/ebs/pubs/wpapers/2008/wp6-08.pdf
- Kandil, M. S., El-Debeiky, S. M. & Hasanién, N. E. (2001). Overview and comparison of long-term forecasting techniques for a fast developing utility: Part I. *Electric Power Systems Research* 58, 11-17
- Kermanshahi, B. S., Poskar, C. H., Swift, G., KcLaren, P., Pedrycs, W., Silk, A. & Buhr, W. (1993). Load forecasting under extreme climatic conditions. *IEEE/NTUA Athens Power Tech Conference: "Planning, Operation and Control of Today's Electric Power Systems"*, Athens, Greece, Sept. 55.
- Leadbetter, M. R. (1983). Extremes and local dependence in stationary sequences. *Probability Theory and Related Fields*, Vol. 65, No. 2, 291-306.

- Loredo, E. N. (2001). *Annual Electrical Peak Load Forecasting Methods with Measures of Prediction Error*. Tesis de Doctorado, Arizona State University.
- Lotufo, A. D. P. & Minussi, C. R. (1999). Electric power systems load forecasting: A survey. *IEEE Power Tech' 99 Conference*, paper BPT99-028-25, Budapest, Hungary.
- Menge, E. E., Wilreker, V. E. & Northcote-Green, J. E. D. (1977). Electrical loads can be forecasted for distribution planning. *Proceedings of American Power Conference*, Vol 39, No. 39, Annual Meeting of the American Power Conference, April 18, Chicago, IL, USA.
- Mirasgedis, S., Sarafidis, E., Georgopoulou, E., Lalas, D. P., Moschovits, M., Karagiannis, F., & Papakonstantinou, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. *Energy*, No 31, 208-227.
- Montgomery, D. C. (1990). *Forecasting and Time Series Analysis* (2nd ed.), New York: Mc Graw Hill.
- Murto, P. (1998). *Neural Network Models for Short-Term Load Forecasting*. Tesis de Maestría, Helsinki University of Technology. Disponible en: <http://www.sal.hut.fi/Publications/pdf-files/tmur98.pdf>
- Nelson, C. R. (1973). *Applied Time Series Análisis for Managerial Forecasting*. San Francisco: Holden-Day, Inc.
- Reiss, R. D. & Thomas, M. (1997). *Statistical Analysis of Extreme Values.*, Basel, Switzerland: Birkhäuser.
- Schmill, J. V. (1968). *Proyecciones de México desde 1968 al año 2000*. México: Compañía de Luz y Fuerza del Centro.
- Smith, R. L. (1985), Maximum likelihood estimation in a class of nonregular cases. *Biometrika* 72, 67–90.

- Smith, R. L. (2004). Statistics of extremes, with applications in environment, insurance and finance. En Finkenstädt, B. and Rootzén, H. (Eds.), *Extreme Values in Finance, Telecommunications and the Environment*. London, U.K.: Chapman & Hall/CRC.
- Stanton, K. N. (1969). Long range demand forecasting for the electric utility industry. *Power Industry Computer Applications Conf.* Denver, CO, May 18-21, 565-579.
- Stanton, K. N. & Gupta, P. C. (1970). Forecasting annual or seasonal peak demand in electric utility systems. *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, No. 5/6, 951-959.
- Temraz, H. K. & Quintana, V. H. (1992). Analytic spatial electric load forecasting methods: A survey. *Can. J. Elect. & Comp. Eng.*, Vol 17, No. 1.
- Willis, H. L. (2002). *Spatial Electric Load Forecasting* (2^a ed.). New York: Marcel Dekker Inc.
- Willis, H. L. (2004). *Power Distribution Planning Reference Book* (2^a ed.). New York: Marcel Dekker Inc.
- Willis, H. L. & Northcote-Green, J. E. D. (1982). A hierarchical recursive method for substantially improving trending of small area load forecasts. *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-101, No. 6.
- Willis, H. L. & Northcote-Green, J. E. D. (1984). Comparison test of fourteen distribution load forecasting methods. *IEEE Transactions on Power Apparatus and Systems*, Vol PAS-103, No. 6, 1190-1197.
- Willis, H. L., Powell, R. W. & Wall, D. L. (1984). Load transfer coupling regression curve fitting for distribution load forecasting. *IEEE Transactions on Power Apparatus and Systems*, Vol PAS-103, No. 5, 1070-1076.
- Willis, R. E. (1979). Statistical considerations in the fitting of growth curves. *Technological Forecasting and Social Change*, No 15, 107-125.