

**HACIA EL INVOLUCRAMIENTO
DE LA ESTADISTICA
EN LA COMPRESION DE LOS PROBLEMAS
DEL MEDIO AMBIENTE.**

118

Francisco Mirabal García Miguel Nakamura Savoy
Centro de Investigación en Matemáticas, A.C.

C I M A T

Apdo. Postal 402

36000-Guanajuato, Gto. México

Tel. (471) 271-55

Fax(471) 257-49

118

Junio de 1992

HACIA EL INVOLUCRAMIENTO DE LA ESTADISTICA EN LA COMPRESION DE LOS PROBLEMAS DEL MEDIO AMBIENTE.

Francisco Mirabal García Miguel Nakamura Savoy
Centro de Investigación en Matemáticas, A.C.

C I M A T
Apdo. Postal 402
36000-Guanajuato, Gto. México
Tel. (471) 271-55
Fax(471) 257-49

Introducción

Es de todos sabido que la diversidad de actividades del ser humano relativas al desarrollo industrial, tecnológico, de transformación, agrícolas, entre otras, han generado contaminantes de muchas clases cuyo impacto se manifiesta en el deterioro del medio ambiente.

La contaminación en el aire producida por industrias de muy diversa índole, por plantas de energía, la debida a exhaustivas emisiones de vehículos automotores, los pesticidas así como detergentes, refrigerantes o contaminación ambiental por aguas negras, son algunos de los problemas que enfrenta la población mundial. Generalmente, la contaminación atmosférica es un fenómeno que se presenta en ciudades con una alta concentración de población, de industrias y transporte.

Hasta ahora, los estudios de monitoreo e investigaciones realizadas en nuestro país sobre contaminación generalmente se han orientado para cuantificar e identificar la cantidad de contaminantes que se hallan en el medio ambiente, así como también para establecer cierto tipo de tendencias en el comportamiento de los contaminantes y para detectar problemas potenciales. Otros estudios buscan determinar como se distribuye la contaminación y como persiste en el aire, agua, suelo o

subsuelo y en la vegetación y de que manera afecta al hombre y su entorno.

Sin embargo, falta profundizar en el estudio e investigación de fuentes naturales de contaminación producto de amplias extensiones de tierras cultivables erosionadas o deforestadas, de filtraciones a los mantos acuíferos de residuos tóxicos por confinamiento inadecuado de los mismos, así como por aguas industriales no tratadas. Otros ejemplos serían fuentes artificiales de contaminación por procesos industriales de fundición o químicos, por plantas generadoras de energía eléctrica, industrias de cemento o de fertilizantes, todas ellas productoras de partículas suspendidas asociadas con otros agentes contaminantes como residuos líquidos, sólidos o gaseosos.

También, es importante profundizar en estudios para establecer si el problema de la contaminación ha tenido alguna declinación como resultado de programas o medidas de control, su eficacia y la validación de los mismos. Evaluar los métodos de pronóstico implementados hasta ahora y que a su vez permitan con mayor grado de certidumbre determinar y cuantificar los impactos reales y las relaciones causa-efecto de las decisiones que se toman.

Dentro de este contexto, uno de los grandes retos de nuestro país, es el que presenta la contaminación ambiental en la Ciudad de México. Este, es sin lugar a dudas un problema que exige de un trabajo de investigación multidisciplinario y globalizador; en razón de que contribuyen en él numerosos factores como lo son las condiciones atmosféricas muy particulares del Valle de México, las complejas reacciones entre los

contaminantes atmosféricos primarios: monóxido de carbono(CO), bióxido de azufre (SO₂), bióxido de nitrógeno (NO₂) e hidrocarburos. El ozono fotoquímico (O₃), contaminante de los llamados secundarios, surge precisamente a través de una serie de reacciones fotoquímicas generadas por los hidrocarburos y óxidos de nitrógeno y compuestos orgánicos que se combinan cuando son expuestos a los rayos ultravioleta que provienen de la luz solar. Es importante subrayar que el ozono es el contaminante con una gran presencia, permanencia y tendencia hacia el incremento en la atmósfera de la Cd. de México, como lo han demostrado los niveles alcanzados en los primeros meses del presente año.

Por otra parte, el asentamiento en la zona metropolitana de poco más de 30 000 industrias de índole diversa y los aproximadamente 3.5 millones de vehículos automotores que circulan en dicha zona han agravado de manera considerable el problema de la contaminación del aire en la Cd. de México.

Al igual que para muchas otras disciplinas científicas, el estudio del problema de la contaminación exige a la estadística la aplicación de técnicas de análisis de datos, la aplicación de modelos de probabilidad y metodología estadística especialmente diseñada para poder abordar el problema desde un punto de vista científico y no de especulación.

En este sentido, el Area de Probabilidad y Estadística del Centro de Investigación en Matemáticas A.C. (CIMAT), de la ciudad de Guanajuato, Gto., tiene entre uno de sus proyectos establecidos, un programa de investigación con el objetivo central de fomentar la investigación para el estudio estadístico de la contaminación en México.

De manera más específica, este programa busca realizar investigación sobre modelos estadísticos que coadyuven a comprender, evaluar y predecir la contaminación en el caso de la Cd. de México, sin olvidar el estudio y en su caso la propuesta de modelos estadísticos para otros problemas particulares de contaminación en el país.

En la misma orientación, el Area de Probabilidad y Estadística busca promover el necesario intercambio interdisciplinario entre aquellos investigadores interesados en modelos estadísticos para la contaminación y sus efectos en el entorno.

En este sentido, la I Reunión de Estadística y Contaminación celebrada en el mes de noviembre de 1991 en el Centro de Investigación en Matemáticas A.C. (CIMAT), de la ciudad de Guanajuato, Gto., tuvo como objetivo reunir a investigadores nacionales, quienes presentaron para su discusión y análisis, un amplio espectro de ponencias e investigaciones: estudios sobre la calidad del aire y sus efectos en la vegetación y en la salud, la diseminación de plaguicidas agrícolas, contaminación marina, evaluaciones estadísticas del contenido de nitratos en aguas subterráneas, correlación empírica entre ozono y sus precursores, distribución del ingreso y su relación con la contaminación atmosférica. Se analizaron también la caracterización del papel de la estadística y el medio ambiente, actividades y oportunidades. El evento reunió a más de un centenar de profesionistas de todo el país, entre los que se contaban ingenieros ambientales, químicos, agrónomos, metalúrgicos, biólogos, epidemiólogos, economistas, sociólogos, estadísticos y de otras ramas de

las matemáticas, así como a estudiantes de diversas licenciaturas y maestrías. Además, cabe señalar que para el próximo año, se tiene programada la II Reunión de Estadística y Medio Ambiente, misma que se celebrará en Universidad de Sonora, en mayo de 1993, en la que también se convocarán a investigadores extranjeros.

Una de las características más importantes en relación con el problema de la contaminación atmosférica, es que las fuentes y los niveles de contaminantes, no se encuentran en la gran mayoría de los casos trivialmente distribuidas y por su naturaleza las mediciones de monitoreo ambiental frecuentemente se pierden. En buena parte, ésto dificulta la aplicación de métodos estadísticos tradicionales.

Sin embargo, consideramos que la estadística debe utilizarse como herramienta de predicción y análisis, y que, aún cuando ya se han empleado y aplicado en algunos estudios técnicas de análisis estadístico clásico y de inteligencia artificial que han incluido parámetros de inversión térmica, de superficie y de altura para pronosticar por ejemplo niveles de ozono [4], se hace necesaria una mayor y profunda investigación sobre el fenómeno contaminación a través de modelos más complejos como valores extremos multivariados, estadística multivariada, nuevas técnicas sobre muestreo y graficación de datos, y de regresión no convencional.

Por otra parte, es importante señalar que las investigaciones, así como los desarrollos tecnológicos orientados a atacar el problema de la contaminación en el país son recientes, como también lo son a nivel mundial.

Se tiene un primer avance en el país, en lo realizado por algunos grupos de investigación que están intentando desarrollar y formular modelos basados en métodos numéricos de los procesos atmosféricos ligados con la química de los contaminantes. El objetivo es predecir cambios ambientales que puedan ocurrir ante la presencia de agentes químicos. Se clasifican los modelos relacionados con la contaminación atmosférica, en dos grupos: los modelos tipo pronóstico y los modelos diagnóstico [6]. La formulación de modelos numéricos de simulación relacionados en particular con la calidad del aire, parten de un conjunto de suposiciones y aproximaciones que permiten traducir el problema real en un problema físicamente idealizado.

Los modelos diagnóstico se basan en la descripción histórica de los datos obtenidos sobre calidad del aire, por lo que implícitamente consideran métodos estadísticos convencionales. Los modelos de simulación poseen su propia metodología y sus objetivos, ya que pretenden generar escenarios de contaminación a partir de un escenario base dictado por condiciones iniciales, como pueden ser por ejemplo, los campos de vientos, las temperaturas de la atmósfera, la radiación solar y las cantidades de emisiones de sustancias contaminantes y/o precursoras.

El desarrollo del escenario a lo largo de tiempo, es decir, el desplazamiento, formación y dispersión o destrucción de contaminantes, se simula obedeciendo leyes físicas y químicas y con ciertas componentes aleatorias. Al examinar el modelo después de un tiempo transcurrido predeterminado, se observa el estado actual del sistema y se interpreta éste como la predicción del escenario para el tiempo transcurrido, sin

embargo, no se tiene una cuantificación acerca de la certidumbre de dicha predicción. En síntesis, los modelos de simulación, cuando funcionan, permiten en principio evaluar los efectos de alguna política de control, simplemente cambiando las condiciones del escenario inicial.

Los modelos estadísticos, por otra parte, consisten en cuantificar la incertidumbre acerca de un fenómeno aleatorio mediante la obtención de una distribución de probabilidad, basándose en observaciones pasadas del fenómeno. Una solución estadística al problema del ozono consistiría en especificar la probabilidad de que el nivel de este contaminante exceda ciertos límites; esta cuantificación auxiliaría en la toma de decisiones sobre medidas de control o contingencia.

En este sentido, se encuentra en proceso de investigación y desarrollo en el CIMAT, un modelo estadístico para niveles de ozono en la Cd. de México, empleando como covariables situaciones meteorológicas [3], basado en un análisis estadístico de los datos de ozono del sistema de monitoreo atmosférico en la zona metropolitana del Valle de México [5].

Generalidades Sobre Un Modelo Estadístico Para Excedentes de Ozono en la Cd. de México. Un Ejemplo.

Ha manera de ilustración, presentaremos a continuación una aplicación de técnicas estadísticas a un problema específico: el de los excedentes de los niveles de ozono en la Cd. de México. Este trabajo ha sido realizado por Miguel Nakamura, Victor Pérez-Abreu, investigadores del CIMAT y Rafael Pérez-Abreu de Teléfonos de México. Los resultados preliminares se

reportan en [3]. Aquí únicamente se mencionarán algunas generalidades no-técnicas. Pretendemos con este ejemplo ilustrar tres ideas fundamentales: primero, el que la estadística cuantifica la aleatoriedad de algún fenómeno a través de su distribución de probabilidad; segundo, que algunas peculiaridades de los datos ambientales hacen necesarias algunas técnicas de desarrollos más recientes y finalmente, el que la estadística necesariamente interactúa con otras disciplinas para lograr soluciones a problemas complejos en los cuales la aleatoriedad juega un papel preponderante.

La Red Automática de Monitoreo Atmosférico (RAMA) de la zona metropolitana, realiza lecturas de concentraciones de ozono (y otros contaminantes) en partes por millón (ppm) cada minuto, con las cuales se obtiene un promedio por hora. En lo que sigue, la observación básica considerada será el máximo diario (Anexo fig. 1) [5].

Debido a fallas en equipos de medición y otros factores, es común encontrar gran cantidad de *observaciones perdidas* en datos de monitoreo del medio ambiente. Cabe destacar que esta no es una característica de México sino del mundo entero. En general, las observaciones perdidas y la no información constituyen una de las características que frecuentemente poseen los datos ambientales. Otras que impiden la aplicación directa de métodos estadísticos elementales son por ejemplo, la *dependencia* entre dos o más mediciones, la ausencia subyacente de una distribución de probabilidad convencional como la normal o alguna distribución simétrica, la frecuente ocurrencia de observaciones por debajo del nivel de detección de los aparatos de medición, y la ocurrencia de *valores extremos*

(observaciones aparentemente aberrantes con respecto al resto de la masa de los datos). Al no cumplirse los supuestos ordinarios de los modelos estadísticos clásicos, se hace inadecuada o imposible la aplicación de métodos tales como el análisis de regresión o el análisis clásico de series de tiempo.

Una alternativa de análisis de difusión más reciente, para el presente ejemplo, es el llamado "enfoque umbral" para el análisis de extremos. Se da el nombre de *Episodio de Ozono* a un grupo de días consecutivos en los que los máximos diarios rebasan un umbral alto determinado de concentración de ozono (Anexo fig. 2) [5]. Recientemente se ha demostrado teóricamente que es posible modelar excedentes de contaminantes mediante la llamada *Distribución de Pareto Generalizada*.

En el caso del ozono, Smith [7] propone un modelo que usa esta distribución, el cual se aplica en Pérez-Abreu [5] a los datos de ozono en cinco estaciones de la Red Automática de Monitoreo Atmosférico (RAMA) de la Cd. de México. Este modelo hace uso de los llamados métodos umbrales, [1], y tiene entre otras ventajas el de integrar de manera especial la falta de información por observaciones perdidas.

Por el momento, sin especificar mayores detalles, bastará mencionar que la *Distribución de Pareto Generalizada* está definida sobre valores positivos y posee dos parámetros: σ (positivo) un parámetro de escala y k , un parámetro de forma. La variación conjunta de ambos parámetros genera una amplia gama de distribuciones que se aplican a observaciones extremas, incluyendo a la distribución exponencial y otras distribuciones

de colas pesadas (Anexo fig. 3) [5].

De acuerdo a los hallazgos de Pérez-Abreu [5], existen razones para suponer que el parámetro de forma (k) es constante sobre todas las estaciones de monitoreo, y que el parámetro de forma (σ) es el que varía de acuerdo a condiciones externas. Esto motiva el suponer que dicho parámetro depende, por ejemplo, de variables meteorológicas. Esto equivale a decir que la distribución de probabilidad de los extremos de ozono depende de las condiciones meteorológicas a través de cierto parámetro. En una fase inicial se consideran las siguientes variables, obtenidas del radio-sondeo de las 6 a.m. realizado en el aeropuerto de la Cd. de México: velocidad del viento a 500 mb, humedad relativa, altura geopotencial, temperatura y tiempo transcurrido para obtener la presión de 500 mb. Adoptando la *Distribución de Pareto Generalizada* y una cierta forma específica para la dependencia del parámetro de escala con las variables explicativas, es posible obtener (empleando métodos numéricos) los estimadores de los parámetros involucrados por el método de máxima verosimilitud, los cuales tendrán propiedades estadísticas óptimas. La técnica mencionada ha sido aplicada a datos obtenidos de la Estación Pedregal [3].

Utilizando ciertas propiedades teóricas, una vez estimados los valores de los parámetros, se ha efectuado un diagnóstico que verifica la validez del modelo. Esto significa que la *Distribución de Pareto Generalizada* es viable, y para unas condiciones meteorológicas dadas es entonces posible obtener los parámetros que describen el comportamiento aleatorio de los extremos de ozono.

Desde el punto de vista estadístico, este resultado preliminar es importante porque significa que el mecanismo aleatorio que gobierna a los extremos de ozono ha sido identificado. Sin embargo, para que el modelo sea útil desde el punto de vista de poder realizar predicciones precisas, es conveniente insertar un mayor número de datos así como incluir otras variables explicativas.

Por tal motivo, se tiene contemplado involucrar en el modelo el nivel de ozono del día anterior, así como otras variables meteorológicas adicionales como el índice de estabilidad de la atmósfera u otras variables sugeridas por consultas estrechas con meteorólogos y otros profesionistas ligados al problema. Por otra parte, es necesario que se involucren al resto de las estaciones de monitoreo, así como información meteorológica a otras presiones adicionales a los 500 mb. Cuando se refine más el modelo será posible así mismo realizar simulaciones de episodios de ozono en cuanto a su duración y su intensidad en base a pronósticos meteorológicos, donde el mecanismo de simulación tomaría en cuenta la *Distribución de Pareto Generalizada* que se ha obtenido estadísticamente con los datos.

Conclusiones

El ecólogo norteamericano Kormondy [2], al expresar algunos conceptos relacionados sobre la Ecología y los problemas inherentes a la contaminación atmosférica, apunta: "*Para muchos de ellos no tenemos conocimientos científicos que permitan hacer predicciones seguras, y aún*

si quisieramos hacerlas careceríamos de datos en que apoyarlas...". A lo anterior podemos responder que la posibilidad, así como la capacidad de generar los conocimientos científicos requeridos para interpretar la complejidad de los fenómenos de contaminación y pronosticar ciertos tipos de comportamiento, se encuentran en proceso. Los estudios permitirán predecir de manera cualitativa y cuantitativa el comportamiento de concentraciones de sustancias y agentes degradantes y contaminantes -como en este caso el ozono-, en áreas urbanas como la Cd. de México u otras entidades semiurbanas o rurales como el campo, ríos, bahías.

Todo lo anterior, será realizable sobre la base del involucramiento de estudios estadísticos modernos apoyados en la interdisciplinariedad, aunados a procesos de predicciones meteorológicas perfeccionadas, y de otras disciplinas que aporten información que permita modelar de manera más eficiente los fenómenos de contaminación.

Precisando, podemos decir que todo lo anterior se podrá lograr de manera más puntual, sólo a través de las interacciones conjuntas de carácter conceptual, metodológico y técnico entre la estadística con la ingeniería química, ambiental, la agronomía, la biología, la medicina, la economía, en suma todas las disciplinas que confluyen en la búsqueda de soluciones al problema de la contaminación y sus efectos en la salud del hombre y en el medio ambiente en general.

Finalmente, aún cuando hemos enfocado la ponencia en un aspecto particular, hemos querido subrayar el papel que juega la estadística como

herramienta de interpretación dinámica de los fenómenos y retos que presenta en la actualidad la contaminación del medio ambiente.

Referencias:

- [1] Davison, C. y R.L. Smith. (1990). *Models for exceedances over high thresholds*. Journal of Royal Statistical Society B. Vol. 52, núm 3, pp 393-442.
- [2] Kormondy, Edward J.(1985). *Conceptos de Ecología*. Alianza Universidad.
- [3] Nakamura, Savoy Miguel., Pérez-Abreu, Carrión R.A., Pérez-Abreu Carrión, V.M. (1992) *Un Modelo Estadístico para Excedentes de Episodios de Ozono en la Ciudad de México*. Comunicaciones del CIMAT.
- [4] Páramo, Víctor H. (1991). *El Pronóstico de SEDUE de Niveles de Ozono en Función de las Condiciones Meteorológicas*. Secretaría de Desarrollo Urbano y Ecología. Ponencia, I Reunión de Estadística y Contaminación., CIMAT.
- [5] Pérez-Abreu, Carrión R.A. (1991). *Análisis Estadístico de los Datos de Ozono del Sistema de Monitoreo Atmosférico en la Zona Metropolitana del Valle de México*. Tesis de Maestría en Estadística. Facultad de Matemáticas, Universidad de Guanajuato.
- [6] Ruíz, Santoyo Ma. Esther G., Cruz, Nuñez Xóchitl. (1991). *Modelos fotoquímicos de contaminantes atmosféricos urbanos*. Revista Ciencia. Vol 42, núm.1, pp 99-109.
- [7] Smith, R.L. (1989). *Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone*. Statistical Science. Vol 4, pp 367-393.

ANEXO

DIA TIPICO DE OZONO ESTACION PEDREGAL

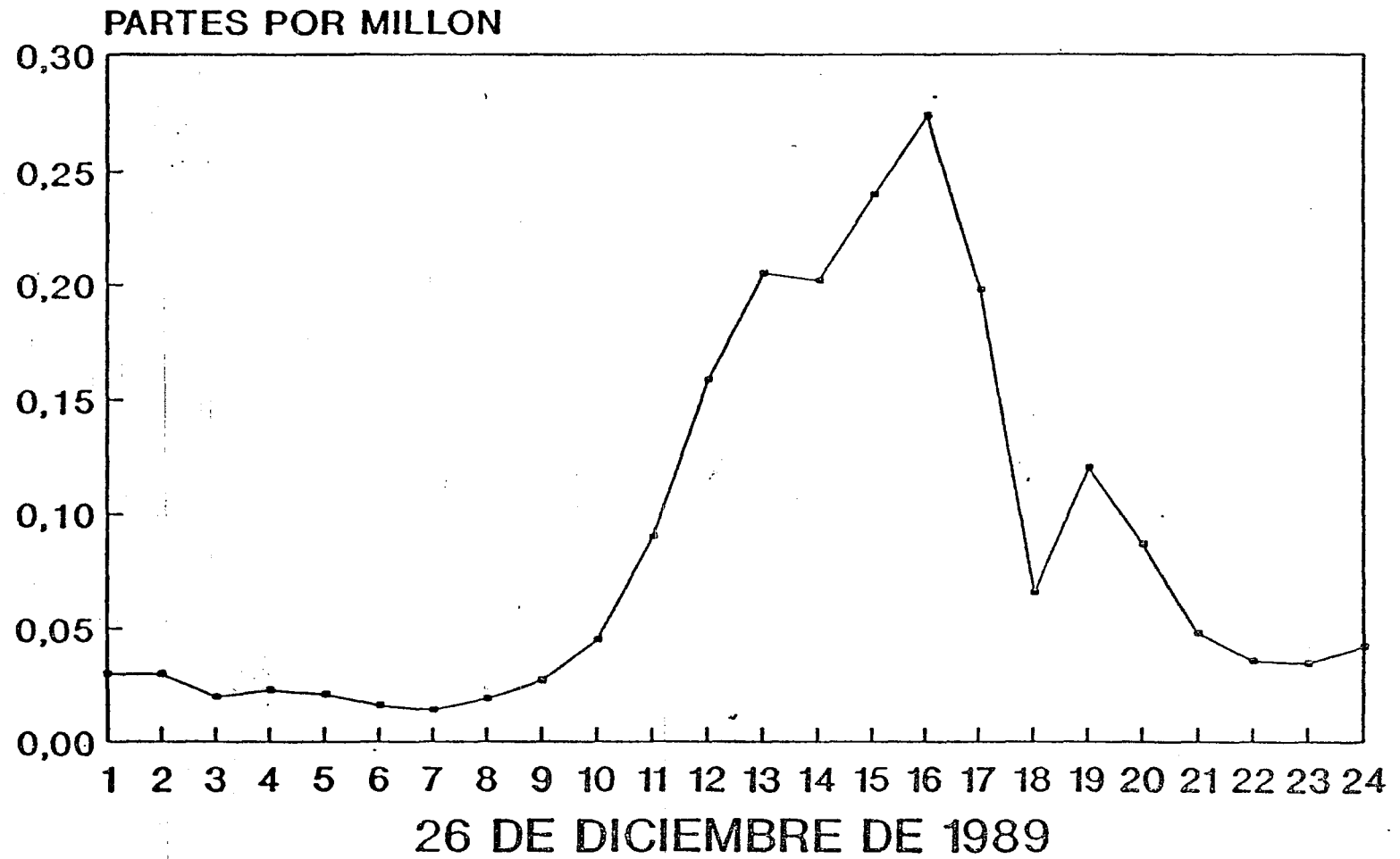


FIGURA 1

EPISODIOS DE OZONO ESTACION PEDREGAL

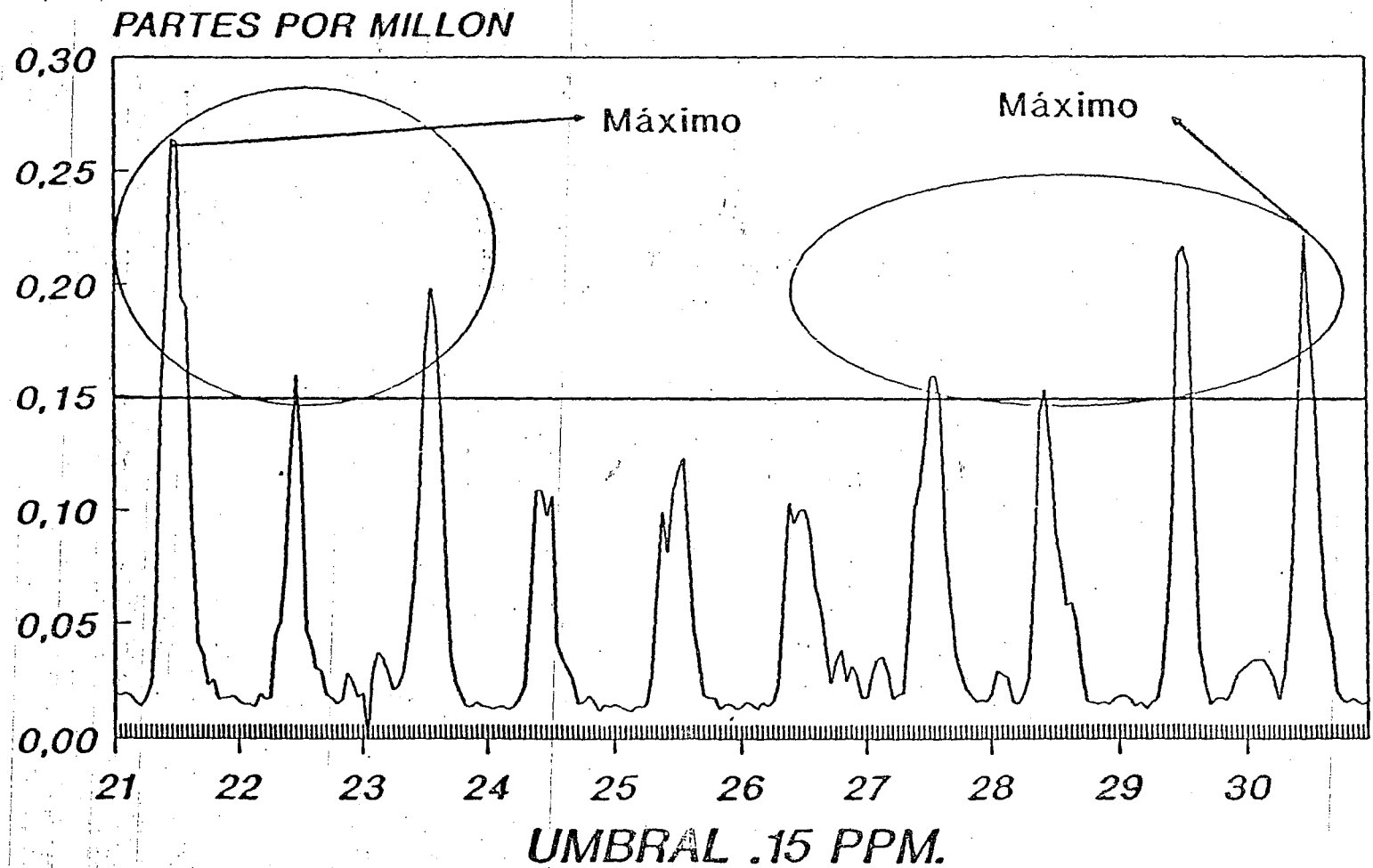


FIGURA 2

DISTRIBUCION PARETO GENERALIZADA DIFERENTES VALORES DE k

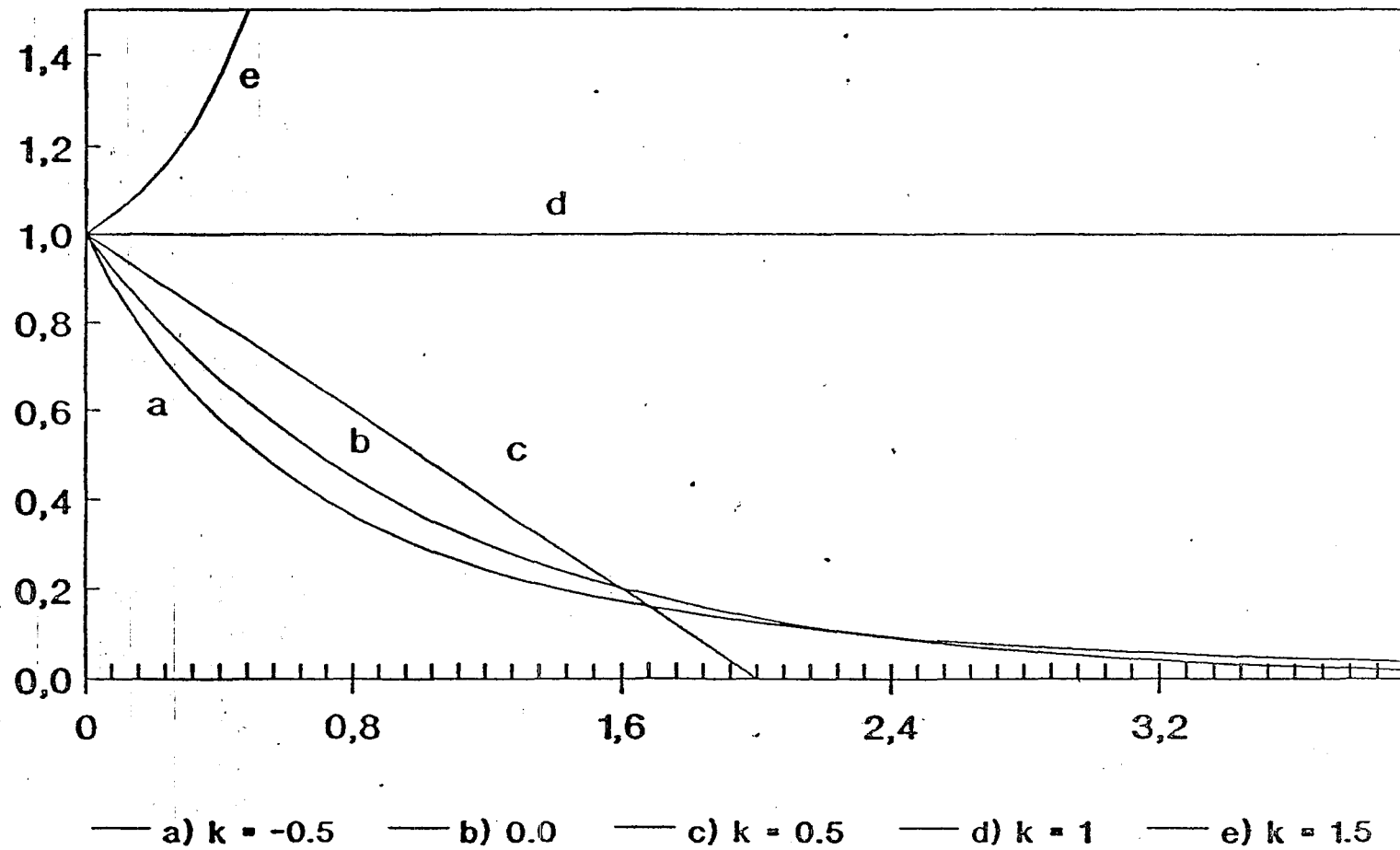


FIGURA 3