

TESIS TESIS TESIS TESIS TESIS



**UNIVERSIDAD AUTÓNOMA  
DE AGUASCALIENTES**

**CENTRO DE CIENCIAS BÁSICAS  
DEPARTAMENTO DE ESTADÍSTICA**

**Estudio Longitudinal de la Contaminación  
Ambiental en el Estado de Guanajuato**

**T E S I S**

que para obtener el grado de

**Maestría en Ciencias**

Área Estadística

**P R E S E N T A :**

**Ing. Liliana Zavala Bayona**

**DIRECTORES DE TESIS:**

M. en C. José de Jesús Ruíz Gallegos

M. en E. Sergio M. Nava Muñoz

Aguascalientes, Ags., Julio de 2009

TESIS TESIS TESIS TESIS TESIS



Centro de Ciencias Básicas

**ING. LILIANA ZAVALA BAYONA**  
**ALUMNO (A) DE LA MAESTRÍA EN CIENCIAS**  
**EXACTAS, SISTEMAS Y DE LA INFORMACIÓN**  
**ÁREA DE ESTADÍSTICA**  
**P R E S E N T E .**

Estimado (a) Sr. (ta) Zavala:

Por medio de este conducto me permito comunicar a Usted que habiendo recibido los votos aprobatorios de los revisores de su trabajo de tesis titulada **“Estudio Longitudinal de la Contaminación Ambiental en el Estado de Guanajuato”**, de acuerdo con su contenido y para dar cumplimiento a los establecido en el Artículo 162 (Fracciones I, II y III) y Artículo 175 (Fracciones I y II) del Reglamento de docencia de la U.A.A., me permito comunicarle que puede imprimir dicho documento y continuar con los trámites para la presentación de su examen de grado.

Sin otro particular me permito saludarle muy afectuosamente.

**A T E N T A M E N T E**  
Aguascalientes, Ags., 30 de junio de 2009  
**“SE LUMEN PROFERRE”**  
**EL DECANO**

**DR. FRANCISCO JAVIER ÁLVAREZ RODRÍGUEZ**



c. c. p.- Archivo

FJAR,mjda





UNIVERSIDAD AUTÓNOMA  
DE AGUASCALIENTES



Asunto: Solicitud de Autorización de  
Examen de Grado de Maestría

DR. FRANCISCO JAVIER ÁLVAREZ RODRÍGUEZ  
DECANO DEL CENTRO DE CIENCIAS BÁSICAS.  
P R E S E N T E.

Por medio de la presente hacemos de su conocimiento que la **Ing. Liliana Zavala Bayona** egresada de la Maestría en Estadística, perteneciente al Posgrado en Ciencias Exactas, Sistemas y de la Información, ha presentado el documento final de su tesis de maestría titulado **“Estudio Longitudinal de la Contaminación Ambiental en el Estado de Guanajuato”**.

La tesis reúne todos los requisitos académicos exigidos, esto es, los elementos teóricos, metodológicos y técnicos, que permiten que la investigación sea defendida en el examen de grado reglamentario.

Por lo anterior, se solicita su autorización para que la Ing. Zavala proceda a los trámites correspondientes para la presentación del examen de grado de maestría.

Atentamente  
“Se Lumen Proferre”

Aguascalientes, Ags., a 29 de Junio de 2009

M. en C. José de Jesús Ruíz Gallegos

Tutor

M. en E. Sergio Martín Nava Muñoz

Tutor

M. en C. Silvia Rodríguez Narciso  
Sinodal



## Dedicatorias

A mis padres, que me dieron la vida, que me formaron con su ejemplo, que se entregaron sin reservas a mi educación, a ellos que día a día me impulsaron con su aliento y con su fe a seguir adelante. A ellos a quien todo les debo.

A mi familia, por su cariño, apoyo y consejo oportuno siempre que los he necesitado.

A mi prometido, por su paciencia y apoyo en todo momento. Por su entrega y por su gran amor. Por lo que somos y por lo que llegaremos a ser juntos.

A mis compañeras y además amigas de maestría, porque juntas nos retroalimentamos y crecimos.





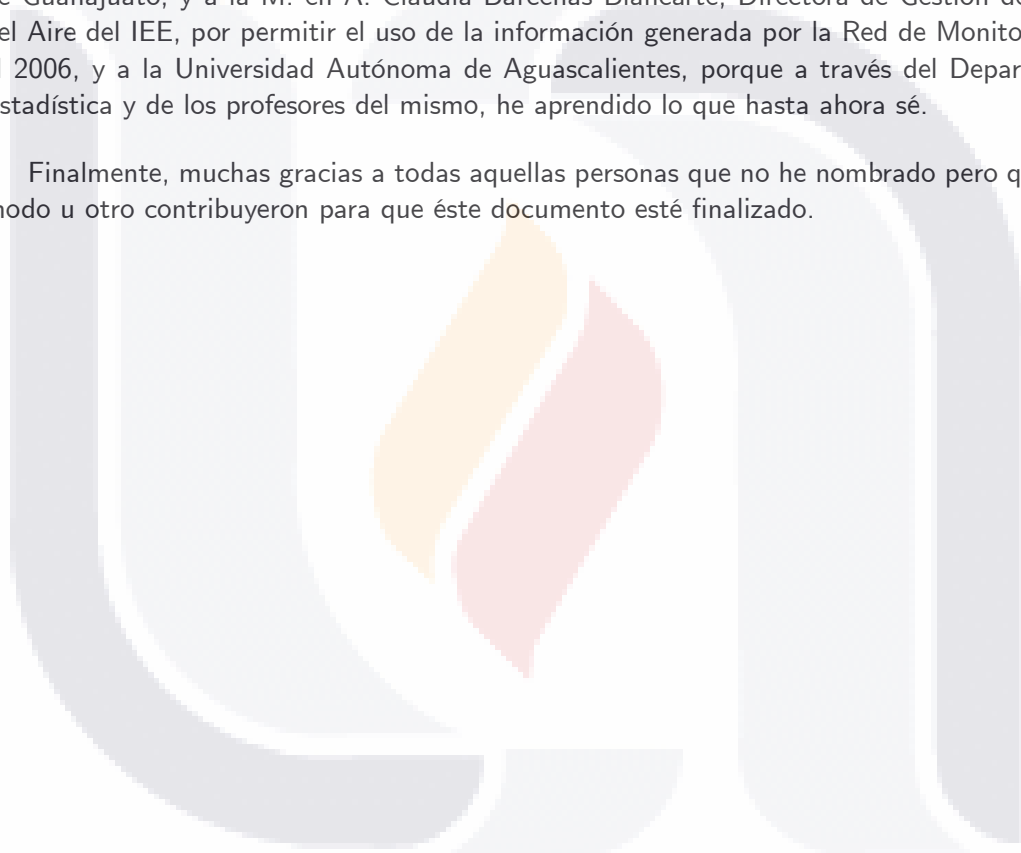


## Agradecimientos

Un sincero agradecimiento a las personas que contribuyeron en la realización de éste esfuerzo, de manera particular expreso una sincera gratitud a las generosas contribuciones de los maestros asesores M. en C. José de Jesús Ruíz Gallegos (U.A.A.), y M. en E. Sergio M. Nava Muñoz (CIMAT Aguascalientes), quienes siempre demostraron una espléndida disposición para guiar, orientar y contribuir de manera incondicional, entusiasta y acertada con ésta tesis. También agradezco a M. en C. Silvia Rodríguez Narciso por sus aportaciones a la misma.

De igual manera se extiende el agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACyT), sin cuyo apoyo económico durante el último semestre de la maestría, éste estudio no se hubiera logrado, al M. en C. Enrique Kato Miranda, Director del Instituto de Ecología del Estado de Guanajuato, y a la M. en A. Claudia Bárcenas Blancarte, Directora de Gestión de la Calidad del Aire del IEE, por permitir el uso de la información generada por la Red de Monitoreo durante el 2006, y a la Universidad Autónoma de Aguascalientes, porque a través del Departamento de Estadística y de los profesores del mismo, he aprendido lo que hasta ahora sé.

Finalmente, muchas gracias a todas aquellas personas que no he nombrado pero que de algún modo u otro contribuyeron para que éste documento esté finalizado.

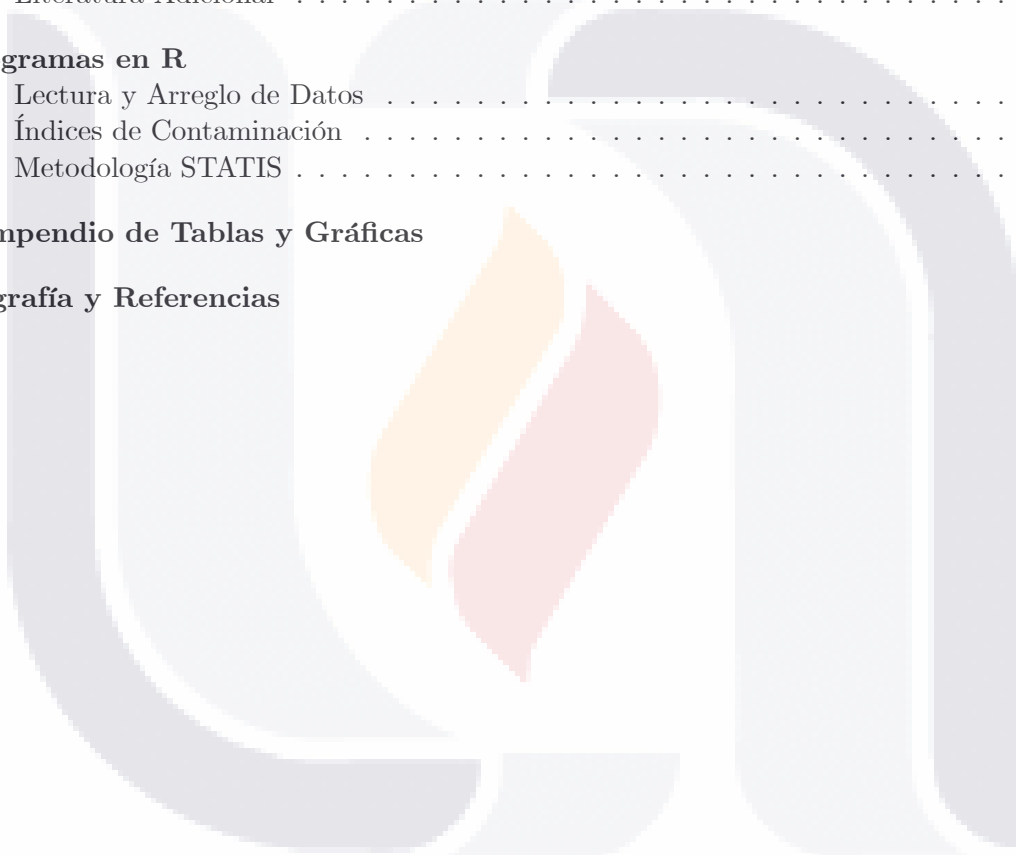




# Índice del Contenido

Resumen	vii
Introducción	1
<b>1 Contaminación Atmosférica en el Corredor Industrial del Bajío</b>	<b>3</b>
1.1 Red de Monitoreo	3
1.1.1 Corredor Industrial del Bajío	5
1.2 Contaminantes	8
1.2.1 Contaminación Atmosférica	8
1.2.2 Fuentes de Contaminación Atmosférica	9
1.3 Contaminantes Criterio	9
1.3.1 Óxidos de Azufre ( $\text{SO}_x$ )	10
1.3.2 Material Particulado (PST, $\text{PM}_{10}$ , $\text{PM}_{2.5}$ )	12
1.3.3 Óxidos de Nitrógeno ( $\text{NO}_x$ )	13
1.3.4 Ozono ( $\text{O}_3$ )	13
1.3.5 Monóxido de Carbono (CO)	14
1.3.6 Normas de Calidad del Aire	15
1.4 Variables Meteorológicas	16
1.4.1 Presión Barométrica	17
1.4.2 Temperatura Ambiente	17
1.4.3 Humedad Relativa	17
1.4.4 Radiación Solar	17
1.4.5 Dirección del Viento	18
1.4.6 Velocidad del Viento	18
1.4.7 Sensores para medición de variables meteorológicas	19
1.5 Validación de la Información	19
1.5.1 Transferencia de datos de las estaciones al Centro de Control	19
1.5.2 Validación de Datos	20
<b>2 Metodología STATIS</b>	<b>21</b>
2.1 Análisis de Componentes Principales	21
2.2 Metodología STATIS	22
2.2.1 Notación	24
2.2.2 Interestructura	24
2.2.3 El Compromiso	24
2.2.4 Intraestructura	24
2.2.5 Estandarización de los Datos	25

2.2.6	Matriz de Productos Escalares . . . . .	25
<b>3</b>	<b>Aplicación</b>	<b>27</b>
3.1	Imputación de Información . . . . .	27
3.2	Análisis de Componentes Principales de la matriz de Correlaciones . . . . .	31
3.3	Intraestructura e Interestructura . . . . .	47
3.4	Compromiso . . . . .	49
3.5	Trayectorias . . . . .	52
	<b>Conclusiones</b>	<b>57</b>
<b>A</b>	<b>Análisis de Datos y Gráficos en R</b>	<b>59</b>
A.1	Características . . . . .	59
A.2	Literatura Adicional . . . . .	60
<b>B</b>	<b>Programas en R</b>	<b>61</b>
B.1	Lectura y Arreglo de Datos . . . . .	61
B.2	Índices de Contaminación . . . . .	65
B.3	Metodología STATIS . . . . .	67
<b>C</b>	<b>Compendio de Tablas y Gráficas</b>	<b>73</b>
	<b>Bibliografía y Referencias</b>	<b>109</b>



# Índice de Figuras

1.1	Logotipo de la Red de Monitoreo . . . . .	4
1.2	Distribución de las Estaciones de Monitoreo por Ciudad . . . . .	4
1.3	Red de Monitoreo de 1999 a 2006 . . . . .	5
1.4	Estaciones de Monitoreo en Salamanca . . . . .	6
1.5	Estación de Monitoreo en Silao . . . . .	6
1.6	Estaciones de Monitoreo en Celaya . . . . .	7
1.7	Estaciones de Monitoreo en León . . . . .	7
1.8	Estaciones de Monitoreo en Irapuato . . . . .	8
1.9	Clasificación General de los Contaminantes . . . . .	9
1.10	Contaminantes Criterio y Variables Meteorológicas . . . . .	10
1.11	Parámetros monitoreados y sus unidades correspondientes . . . . .	11
1.12	Molécula de Dióxido de Azufre . . . . .	11
1.13	Molécula de Dióxido de Nitrógeno . . . . .	13
1.14	Molécula de Ozono . . . . .	14
1.15	Molécula de Monóxido de Carbono . . . . .	14
1.16	Instrumentos de Medición de Contaminantes Criterio . . . . .	15
1.17	Instrumentos de Medición de Partículas $PM_{10}$ y $PM_{2.5}$ . . . . .	15
1.18	Normas de Calidad del Aire . . . . .	16
1.19	Rosa de los Vientos . . . . .	18
1.20	Sensores Meteorológicos . . . . .	19
3.1	Descripción de las variables de la Base de Datos (1) . . . . .	28
3.2	Descripción de las variables de la Base de Datos (2) . . . . .	29
3.3	Porcentaje de datos correctos en la estación de Bomberos de Irapuato . . . . .	30
3.4	Porcentaje de datos correctos en la estación de la Secundaria Oficial de Irapuato . . . . .	30
3.5	Porcentaje de datos correctos en la estación de la primaria Teódula Fuerte Orozco de Irapuato . . . . .	30
3.6	Caja y Brazos de la velocidad del viento en la ciudad de León . . . . .	31
3.7	Histograma de la velocidad del viento en la ciudad de León . . . . .	32
3.8	Gráfica de la velocidad del viento en la ciudad de León . . . . .	32
3.9	Valores Máximos de las Variables en el mes de Enero . . . . .	33
3.10	Valores Máximos de las Variables en el mes de Febrero . . . . .	33
3.11	Valores Máximos de las Variables en el mes de Marzo . . . . .	33
3.12	Datos Estandarizados del mes de Enero . . . . .	33
3.13	Datos Estandarizados del mes de Febrero . . . . .	33
3.14	Matriz de Correlaciones de las Variables en Enero . . . . .	34
3.15	Matriz de Correlaciones de las Variables en Febrero . . . . .	34

3.16	Eigenvalores de R y su Pocentaje de Variación para cada mes del año . . . . .	35
3.17	Gráficas de Laderas de Enero a Abril de 2006 . . . . .	36
3.18	Gráficas de Laderas de Mayo a Agosto de 2006 . . . . .	36
3.19	Gráficas de Laderas de Septiembre a Diciembre de 2006 . . . . .	37
3.20	Puntajes del Primer Componente Principal por mes . . . . .	38
3.21	Puntajes del Segundo Componente Principal por mes . . . . .	39
3.22	Puntajes de las Ciudades de Guanajuato en Enero . . . . .	39
3.23	Puntajes de las Ciudades de Guanajuato en Febrero . . . . .	40
3.24	Puntajes de las Ciudades de Guanajuato en Marzo . . . . .	40
3.25	Puntajes Mensuales en Celaya . . . . .	41
3.26	Puntajes Mensuales en Irapuato . . . . .	41
3.27	Puntajes Mensuales en Salamanca . . . . .	42
3.28	Puntajes Mensuales en León . . . . .	42
3.29	Correlaciones entre Variables en Enero . . . . .	43
3.30	Correlaciones entre Variables en Febreo . . . . .	43
3.31	Correlaciones entre Variables en Marzo . . . . .	44
3.32	Correlaciones del Ozono a través de los meses . . . . .	44
3.33	Correlaciones del Monóxido de Carbono a través de los meses . . . . .	45
3.34	Correlaciones del Dióxido de Azufre a través de los meses . . . . .	45
3.35	Diferencia en los puntajes de Enero y Febrero de 2006 . . . . .	46
3.36	Diferencia en los puntajes de Enero y Marzo de 2006 . . . . .	46
3.37	Diferencia en los puntajes de Enero y Abril de 2006 . . . . .	46
3.38	Matriz de Productos Escalares de Enero . . . . .	47
3.39	Matriz de Productos Escalares de Febrero . . . . .	47
3.40	Matriz de Productos Escalares de Marzo . . . . .	47
3.41	Intraestructura de Enero de 2006 . . . . .	48
3.42	Intraestructura de Febrero de 2006 . . . . .	48
3.43	Intraestructura de Marzo de 2006 . . . . .	48
3.44	Normas $W_k$ . . . . .	48
3.45	Productos Escalares Normados $RV$ . . . . .	49
3.46	Imagen Euclidiana . . . . .	50
3.47	Compromiso de las Cuidades de Guanajuato . . . . .	50
3.48	Correlación del eje principal con el Compromiso . . . . .	51
3.49	Correlaciones del eje secundario con el Compromiso . . . . .	51
3.50	Correlación de las Variables con el Compromiso . . . . .	52
3.51	Eigenvalores de Sdelta . . . . .	53
3.52	Gráfica de Ladera de Sdelta . . . . .	53
3.53	Trayectorias de las ciudades en el mes de Enero de 2006 . . . . .	54
3.54	Trayectorias de las ciudades de Enero a Febrero de 2006 . . . . .	55
3.55	Trayectorias de las Ciudades de Guanajuato . . . . .	55
A.1	Logotipo de R . . . . .	59

## Índice de Tablas

3.1	Primer Componente Principal de cada mes . . . . .	38
3.2	Segundo Componente Principal de cada mes . . . . .	38







# TESIS TESIS TESIS TESIS TESIS

## Resumen

La zona del Corredor Industrial del Estado de Guanajuato está conformada por los municipios de Celaya, Salamanca, Irapuato, Silao y León, y ha sido tipificada como crítica con respecto a contaminación ambiental desde 1994 por la Norma Oficial Mexicana NOM-085-SEMARNAT.

La Red de Monitoreo de la Calidad del Aire del Estado de Guanajuato comenzó a funcionar en el año de 1999, cuando se instaló la primera estación de monitoreo en el municipio de Salamanca. Al 31 de diciembre del año 2006 la red contaba con 12 estaciones fijas y 1 unidad móvil para el monitoreo. Uno de los retos más importantes para la Red de Monitoreo es conocer el nivel de contaminantes criterio al que está expuesta la población e informar oportunamente sobre la calidad del aire, así como contar con elementos para estimar los efectos en la población y el medio ambiente.

El presente trabajo propone un análisis del comportamiento de la información obtenida durante el 2006 para 6 contaminantes ambientales y 5 variables meteorológicas, en función de los 12 meses del año, basado en la metodología STATIS (*Structuration des Tableaux A Trois Indices de la Statistique*), desarrollada por Christine Lavit en 1986, la cual tiene como objetivos encontrar una estructura común dentro de los arreglos de datos (intraestructura), y encontrar la relación existente entre arreglos (interestructura), para así conocer la evolución de la contaminación a lo largo del tiempo.

Con la información horaria registrada en la red de monitoreo, se obtiene un arreglo de datos de tres entradas con el máximo de cada variable como estadístico. A partir de ello, se aplica un Análisis de Componentes Principales y se obtienen los puntajes de los primeros dos componentes, de los cuales, se utiliza el primero, debido a que es una medida relativa que describe la mayor variación de la información en un momento dado. Este problema se explica con mayor detalle en el Capítulo 3.

En el Capítulo 1 de este trabajo se presenta toda la información referente a los contaminantes ambientales y a las variables meteorológicas medidas durante el 2006. En el Capítulo 2 se describe el Análisis de Componentes Principales y la Metodología STATIS. En el Capítulo 3 se detalla la forma en que se depuró la información recibida por la red de monitoreo del Instituto de Ecología del Estado de Guanajuato y se describe la aplicación de la Metodología STATIS a los datos.

TESIS TESIS TESIS TESIS TESIS



TESIS TESIS TESIS TESIS TESIS

# TESIS TESIS TESIS TESIS TESIS

## Introducción

El Gobierno del Estado de Guanajuato, a través del Instituto de Ecología, es el único organismo que monitorea permanentemente y de manera confiable la calidad del aire en el estado de Guanajuato [CAG2009]. Ésta actividad se realiza a través de la Red de Monitoreo de la Calidad del Aire del Estado de Guanajuato, que funciona los 365 días del año, las 24 horas (excepto cuando recibe mantenimiento). El instituto vigila responsablemente y bajo parámetros internacionales la calidad del aire que se respira, además, toda la información que procesa la Red de Monitoreo es pública.

Actualmente la Red opera a través de 14 estaciones de monitoreo automático y 17 equipos de monitoreo manual distribuidos en los 11 municipios del estado con mayor actividad. Ésta red trabaja con la tecnología más avanzada y está sujeta a un programa anual de verificación y calibración que garantiza el cumplimiento de estándares internacionales. Además, cuenta con un programa anual de mantenimiento preventivo realizado por personal externo a Gobierno del Estado y certificado por el fabricante de los equipos. En el 2007 comenzó a implementarse el sistema de calidad ISO 9001, con la finalidad de fortalecer los procesos y procedimientos para la obtención de los datos generados.

En el 2007 el Instituto de Ecología del Estado de Guanajuato presentó el documento *La Calidad del Aire en Guanajuato. Informe 2006 de estado y de tendencias* [LAR2007]. Este informe presenta una descripción del estado que guarda la Red de Monitoreo de la Calidad del Aire y sus principales retos; también describe el procedimiento realizado para el análisis del estado y tendencias de los contaminantes monitoreados en cada una de las estaciones de monitoreo atmosférico, todo ello realizando un análisis detallado por cada una de las redes locales de monitoreo y plasmando conclusiones por cada contaminante monitoreado en la red estatal.

Este trabajo pretende conocer cómo a través de un estudio longitudinal, se comportan los contaminantes ambientales y las variables meteorológicas en cada una de las ciudades del corredor industrial de Guanajuato.

En el Capítulo 1 se introduce al tema de la contaminación ambiental, donde se explica cada uno de los contaminantes ambientales y cada variable meteorológica. Además, se indican las Normas de Calidad del Aire que establecen las concentraciones máximas permisibles de los contaminantes durante un período definido.

La metodología STATIS parece ser la adecuada para este propósito, por lo cual, las principales características de la metodología se describen en el Capítulo 2.

En el Capítulo 3 se presenta la aplicación del modelo STATIS a un arreglo de datos de 3 entradas. Tanto el análisis de componentes principales, como el de la metodología, se desarrolló en el software R.

Finalmente, en el Apéndice se presentan las bondades del programa R y el código del programa que genera las salidas y gráficas de éste estudio..

TESIS TESIS TESIS TESIS TESIS



TESIS TESIS TESIS TESIS TESIS

## Capítulo 1

# Contaminación Atmosférica en el Corredor Industrial del Bajío

Durante el año 2006 se recolectó la información de indicadores (contaminantes y variables meteorológicas), medidos en el corredor Industrial del Bajío.

El esquema de operación de la red de monitoreo continúa siendo a través de Asociaciones Civiles y de instituciones educativas. Actualmente las redes de Salamanca y Celaya están a cargo del Patronato para el Monitoreo de la Calidad del Aire de Salamanca, A.C.; la Red de Irapuato, es operada por el Patronato para el Monitoreo de la Calidad del Aire de Irapuato, A.C., y las Redes de León y Silao, están a cargo de la Universidad Tecnológica de León (UTL) [LAR2007].

En el proceso de mejora continua y con la finalidad de garantizar una mayor confiabilidad en la información, se han fortalecido los controles operativos y mejorado el equipamiento del software para lograr una mejor sistematización del proceso de validación de información. En este proceso se ha contado con el invaluable apoyo y asesoramiento del Instituto Nacional de Ecología (INE) a través del Centro Nacional de Investigación y Capacitación Ambiental (CENICA), así como del Sistema de Monitoreo Atmosférico del Distrito Federal (SIMAT), quienes han contribuido al crecimiento de la Red.

Los retos por enfrentar son, entre otros, la consolidación del Centro de Control para alcanzar una mejor transmisión de registros en Internet, lo cual implica un desarrollo de capacidades internas; el mejoramiento de la plataforma informática y la consolidación de los procesos de validación y manejo de datos, la evaluación permanente del desempeño a través de la instrumentación de auditorías periódicas, la conformación y consolidación de un laboratorio de estándares, y el desarrollo de capacidades locales para el mantenimiento preventivo y correctivo.

A continuación se da una pequeña explicación de donde se encuentran las estaciones que realizaron las mediciones, así como la forma en que se llevaron a cabo.

### 1.1 Red de Monitoreo

Actualmente la Red de Monitoreo de Calidad del Aire del Estado de Guanajuato cuyo logotipo se muestra en la figura 1.1, opera a través de 14 estaciones de monitoreo automático y 17 equipos de monitoreo manual distribuidos en los 11 municipios del estado con mayor actividad.

Las estaciones de monitoreo automático están concentradas en los municipios del corredor industrial del bajío: Celaya, Salamanca, Irapuato, Silao y León, ver figura 1.2.

En 1999 se instaló la primera estación automática en Salamanca y al 31 de diciembre de 2006,



Figura 1.1: Logotipo de la Red de Monitoreo

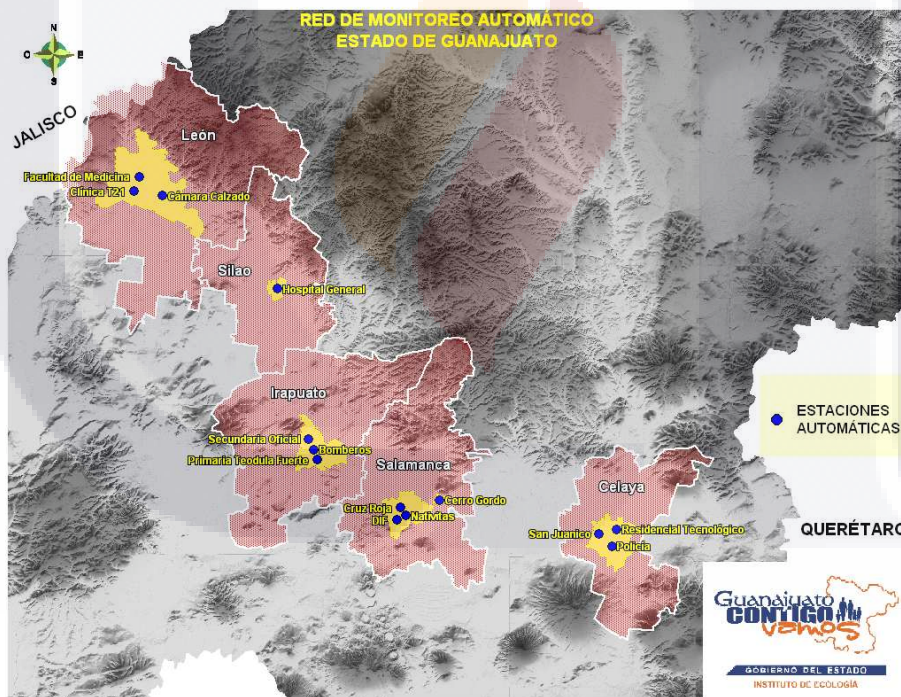


Figura 1.2: Distribución de las Estaciones de Monitoreo por Ciudad

Ciudad	1999	2000	2001	2002	2003	2006	Referencia
Salamanca	Cruz Roja						CR
Salamanca		DIF					DIF
Irapuato		Bomberos					BM
Celaya			San Juanico				SJ
Irapuato			Teodula				TL
Salamanca				Nativitas			NAT
Celaya				R. Tecnológico			RT
Irapuato				Secundaria			SEC
Celaya				Policia mpal.			POL
León					CICEG		CG
Salamanca					Unidad Móvil		UM-1
León						IMSS T21	T21
Silao						Hospital general	SL
<b>Total</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>4</b>	<b>2</b>	<b>2</b>	<b>13</b>

Figura 1.3: Red de Monitoreo de 1999 a 2006

la red contaba con 12 estaciones fijas y 1 unidad móvil para el monitoreo automático. En el 2007 entró en operación la tercera estación en León.

La figura 1.3 muestra la distribución hasta el 2006 de las estaciones de monitoreo automático en cada uno de los 5 municipios del corredor industrial. De las 13 estaciones de monitoreo de dicha tabla, la de León (IMSS T21), la de Salamanca (Unidad Móvil) y la de Silao (Hospital General) no se incluyeron en este estudio. La información de la unidad móvil de Salamanca no era de interés, ya que el estudio solo fue para estaciones fijas. Las casetas ubicadas en Silao y León, se instalaron en el mes de mayo de 2006, por lo que la información mensual estaba incompleta.

El esquema de operación de la red de monitoreo continúa siendo a través de Asociaciones Civiles y de instituciones educativas. Actualmente las redes de Salamanca y Celaya están a cargo del Patronato para el Monitoreo de la Calidad del Aire de Salamanca, A.C.; la Red de Irapuato, es operada por el Patronato para el Monitoreo de la Calidad del Aire de Irapuato, A.C., y las Redes de León y Silao, están a cargo de la Universidad Tecnológica de León (UTL).

### 1.1.1 Corredor Industrial del Bajío

En las figuras 1.4 a la 1.8 se muestra una descripción de cada uno de los municipios, donde es de interés, tanto el parque vehicular como la actividad industrial de cada municipio.

Extensión territorial	745.96 km <sup>2</sup>
Ubicación Geográfica	N: 20°34'08.58" O: 101°11'58.37"
Municipios colindantes	Norte: Juventino Rosas Sur: Valle de Santiago Este: Celaya Oeste: Irapuato
Población	223,623 Hab.
Parque vehicular	57,585 Vehiculos
Actividad industrial	Petroquímica, Química, Generación electricidad, Alimenticia

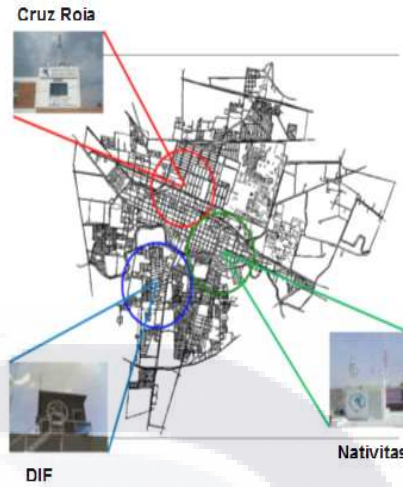


Figura 1.4: Estaciones de Monitoreo en Salamanca

Extensión territorial	531.41 km <sup>2</sup>
Ubicación Geográfica	N: 20°56'37.14" O: 101°25'37.07"
Municipios colindantes	Norte: Guanajuato Sur: Irapuato Este: Guanajuato Oeste: Leon
Población	147123 Hab.
Parque vehicular	21,412 Vehiculos
Actividad industrial	Automotriz

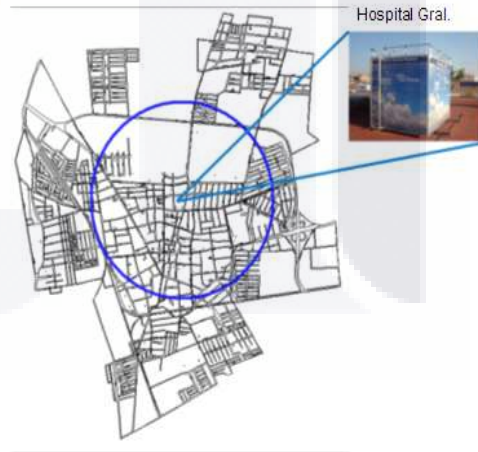


Figura 1.5: Estación de Monitoreo en Silao



Extensión territorial	560.97 km <sup>2</sup>
Ubicación Geográfica	N: 20°31'24.00" O: 100°48'55.00"
Municipios colindantes	Norte: Comonfort Sur: Tarimoro Este: Apaseo el Grande Oeste: Cortazar, Villagran
Población	415,869 Hab.
Parque vehicular	99,418 Vehículos
Actividad industrial	Alimenticia

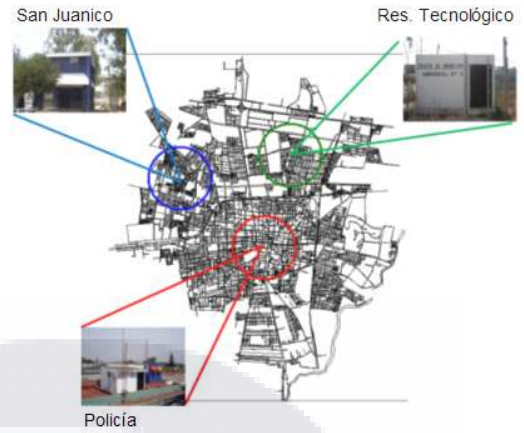


Figura 1.6: Estaciones de Monitoreo en Celaya

Extensión territorial	862 km <sup>2</sup>
Ubicación Geográfica	N: 21°07'14.69" O: 101°40'57.27"
Municipios colindantes	Norte: Lagos Moreno (Jal) Sur: Manuel Doblado, Silao Este: Guanajuato Oeste: San Francisco del Rincón
Población	1,278,087 Hab.
Parque vehicular	281,620 Vehículos.
Actividad industrial	Cuero y Calzado.

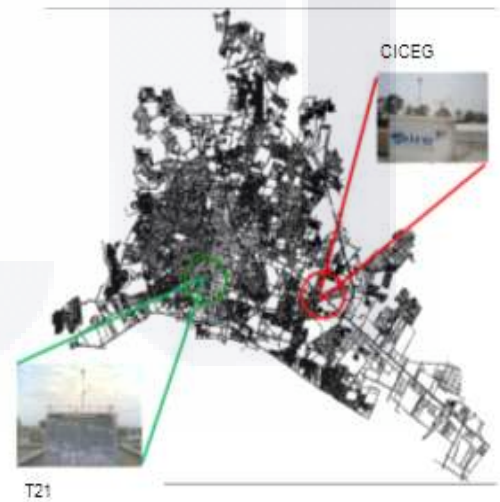


Figura 1.7: Estaciones de Monitoreo en León

Extensión territorial	786,4 km <sup>2</sup>
Ubicación Geográfica	N: 20°40'21.67" O: 101°20'49.70"
Municipios colindantes	Norte: León, Guanajuato Sur: Abasolo Este: Salamanca Oeste: Manuel Doblado
Población	Hab.
Parque vehicular	99,935 Vehículos
Actividad industrial	Alimenticia

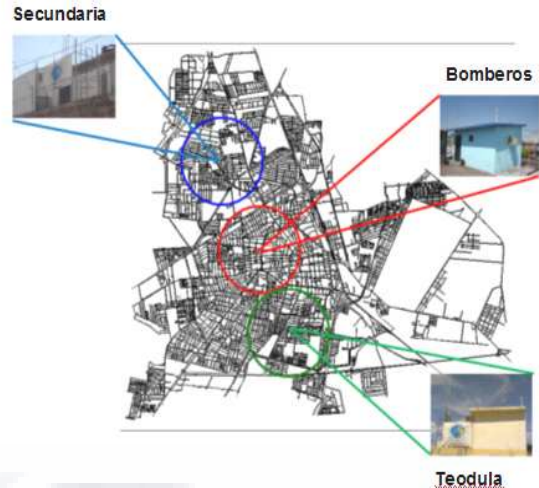


Figura 1.8: Estaciones de Monitoreo en Irapuato

## 1.2 Contaminantes

El aire es una mezcla compleja de sustancias, los principales elementos constitutivos del aire son el nitrógeno, oxígeno y vapor de agua. Aproximadamente 78% del aire es nitrógeno y 21% oxígeno. El uno por ciento restante incluye pequeñas cantidades de sustancias y se expresan en volúmen en partes por millón (ppm) o en partes por billón (ppb): Dióxido de Carbono (350 ppm), Neón (18.2 ppm), Helio (5.24 ppm), Metano (2 ppm), Criptón (1.1 ppm), Hidrógeno (0.5 ppm), Óxido Nitroso (0.3 ppm), Xenón (0.08 ppm), Monóxido de Carbono (0.05-0.2 ppm) y Ozono (0.02-0.03 ppm). También existen partículas sólidas y líquidas en suspensión (por ejemplo humo, polvo, cenizas, sales, materia orgánica, polen, vapores, entre otros).

### 1.2.1 Contaminación Atmosférica

De acuerdo al Artículo 3o. del capítulo I de la Ley General del Equilibrio Ecológico y la Protección al Ambiente (LGEEPA), [LGE2009], se tienen las siguientes definiciones:

Fracción VI. **Contaminación.** Presencia en el ambiente de uno o más contaminantes o de cualquier combinación de ellos que cause desequilibrio ecológico.

Fracción VII. **Contaminante.** Toda materia o energía en cualesquiera de sus estados físicos y formas, que al incorporarse o actuar en la atmósfera, agua, suelo, flora, fauna o cualquier elemento natural, altere o modifique su composición y condición natural.

Fracción VIII. **Contingencia Ambiental.** Situación de riesgo, derivada de actividades humanas o fenómenos naturales, que pueden poner en peligro la integridad de uno o varios ecosistemas.

Se denomina **Contaminación Atmosférica** o contaminación ambiental a la presencia en el ambiente de cualquier agente (físico, químico o biológico) o bien de una combinación de varios agentes en lugares, formas y concentraciones tales que sean o puedan ser nocivos para la salud, la seguridad o para el bienestar de la población; o que puedan ser perjudiciales para la vida vegetal o animal; o impidan el uso normal de las propiedades y lugares de recreación, y el goce de los mismos. La contaminación atmosférica es también la incorporación a los cuerpos receptores de sustancias sólidas, líquidas o gaseosas, o mezclas de ellas, siempre que alteren desfavorablemente



Figura 1.9: Clasificación General de los Contaminantes

las condiciones naturales de los mismos, o que puedan afectar la salud, la higiene o el bienestar del público [DCA2009].

### 1.2.2 Fuentes de Contaminación Atmosférica

Existen 2 fuentes de contaminación:

#### Naturales.

Cuando provienen de la litósfera, hidrósfera o de las plantas o animales. En teoría, el aire siempre ha tenido cierto grado de contaminación debido a fenómenos naturales (erupción de volcanes, tormentas de viento, descomposición de plantas y animales, etc.).

#### Antropogénicas.

Cuando se derivan de cualquier tipo de actividad humana.

### 1.3 Contaminantes Criterio

Varias son las posibles clasificaciones que se pueden realizar al respecto de las sustancias contaminantes del aire que existen. Una de ellas se muestra en la figura 1.9.

#### Contaminantes Criterio.

Son los que se generan por las actividades antropogénicas y que son perjudiciales para la salud y el bienestar de los seres humanos. Desde 1970 la Agencia de Protección Ambiental de los Estados Unidos (*U.S. EPA*) adoptó a éstos como *Contaminantes Criterio* porque previo a su regulación, fueron objeto de estudios científicos sobre la calidad del aire y sentaron los precedentes para establecer los niveles máximos permisibles de los mismos. Los contaminantes criterio son:

Contaminantes criterio		Variables meteorológicas	
SO <sub>2</sub>	Bioxido de azufre	WS	Velocidad de viento
PM <sub>10</sub>	Particulas menores a 10 micras	WD	Dirección de viento
PM <sub>2.5</sub>	Particulas menores a 2.5 micras	T°	Temperatura ambiente
O <sub>3</sub>	Ozono	HR	Humedad relativa
CO	Monoxido de carbono	PB	Presión barométrica
NO <sub>2</sub>	Bioxido de nitrógeno	RS	Radiación solar
*HCT	Hidrocarburos totales	PP	Precipitación pluvial

Figura 1.10: Contaminantes Criterio y Variables Meteorológicas

1. Óxidos de azufre ( $SO_x$ ).
2. Material particulado ( $PM_{10}$ ,  $PM_{2.5}$ ).
3. Óxidos de nitrógeno ( $NO_x$ ).
4. Ozono ( $O_3$ ).
5. Monóxido de carbono ( $CO$ ).

El Instituto de Ecología de Guanajuato mide los contaminantes criterio y las variables metereológicas que se muestran en la figura 1.10.

Las unidades de cada variable se presentan en la figura 1.11, además se indica cuales variables se midieron en cada una de las ciudades (incluida la unidad móvil).

En seguida se describe cada uno de los contaminantes criterio y se indican los efectos en la salud humana.

### 1.3.1 Óxidos de Azufre ( $SO_x$ )

Los óxidos de azufre son gases incoloros, irritantes y tóxicos que se forman al quemar azufre. El dióxido de azufre ( $SO_2$ ) (figura 1.12), es el contaminante criterio que indica la concentración de óxidos de azufre en el aire. Las fuentes primarias de óxidos de azufre son los procesos de combustión, ya que los combustibles como el carbón, el petróleo, el diesel o el gas natural contienen ciertas cantidades de compuestos azufrados.

#### Efectos en la Salud

En altas concentraciones el Dióxido de Azufre ( $SO_2$ ) puede ocasionar dificultad para respirar, humedad excesiva en las mucosas de las conjuntivas, irritación severa en vías respiratorias e incluso al interior de los pulmones por formación de partículas de ácido sulfúrico, ocasionando vulnerabilidad en las defensas. También puede llegar a causar broncoespasmos en personas sensibles como los asmáticos, agravamiento de enfermedades respiratorias y cardiovasculares existentes, y la muerte.

CIUDAD	EST.	O <sub>3</sub>	SO <sub>2</sub>	CO	NO	NO <sub>2</sub>	NOX	PM10	CH <sub>4</sub>	NMHC	THC	TEM	WS	WD	P.BAR	RAD	HR
SALAMANCA	NAT	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>	ppm	ppm	ppm	°C	m/s	°	Milibar	W/m <sup>2</sup>	%
	CR	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>	ppm	ppm	ppm	°C	m/s	°			%
	DIF	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	Milibar	W/m <sup>2</sup>	%
CELAYA	POL	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
	SJ	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
	TEC	ppb			ppb	ppb	ppb					°C	m/s	°	mmHg	W/m <sup>2</sup>	%
IRAPUATO	BOM	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
	TEO	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
	SEC	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg		%
SILAO	SL	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
LEÓN	T21	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
	CG	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>				°C	m/s	°	mmHg	W/m <sup>2</sup>	%
MÓVIL	MV	ppb	ppb	ppm	ppb	ppb	ppb	µg/m <sup>3</sup>	ppm	ppm	ppm	°C	m/s	°	mmHg		%

Figura 1.11: Parámetros monitoreados y sus unidades correspondientes



Figura 1.12: Molécula de Dióxido de Azufre

La combinación de óxidos de azufre y partículas suspendidas actúan sinérgicamente produciendo un efecto combinado mucho más nocivo que el efecto individual de cada uno de ellos por separado. Se ha encontrado que los óxidos de azufre perjudican el sistema respiratorio, especialmente de las personas que sufren asma y bronquitis crónica. Los efectos de los óxidos de azufre empeoran cuando el dióxido de azufre se combina con partículas o humedad del aire.

Los óxidos de azufre también son responsables de algunos efectos sobre el bienestar, el de mayor preocupación es la contribución de óxidos de azufre a la formación de lluvia ácida que puede perjudicar lagos, la vida acuática, materiales de construcción y la vida silvestre.

Si bien los efectos señalados dependen en gran medida de la sensibilidad de cada individuo, los grupos de la población más sensibles al  $SO_2$  incluye a los niños y ancianos, a los asmáticos, y a aquellos con enfermedades pulmonares crónicas como bronquitis y enfisema.

### 1.3.2 Material Particulado ( $PST$ , $PM_{10}$ , $PM_{2.5}$ )

El material particulado se forma por muchos procesos, tales como el viento, polinización de plantas e incendios forestales, aunque su efecto se ve incrementado por actividades humanas que implican el cambio de uso de suelo.

Las principales fuentes antropogénicas de pequeñas partículas incluyen la quema de combustibles sólidos como la madera y el carbón, las actividades agrícolas como la fertilización y la industria de la construcción.

Con la denominación de partículas totales en suspensión ( $PST$ ) se reconoció a una amplia categoría de material particulado como contaminante criterio. Las  $PST$  son las partículas sólidas o líquidas del aire mayores a  $10 \mu m$ , donde se incluyen contaminantes como el polvo, hollín y partículas líquidas producidas por la condensación de vapores.

Se han especificado normas sobre el material particulado con menos de 10 micrómetros de diámetro ( $\mu m$ ), y además al material particulado con menos de 2.5 micrómetros de diámetro. Estas partículas son comúnmente referidas como  $PM_{10}$  y  $PM_{2.5}$ , respectivamente.

#### Efectos en la Salud

Las partículas totales en suspensión ( $PST$ ) son retenidas básicamente en las vías respiratorias superiores y eliminadas en su mayor parte por el sistema de limpieza natural del tracto respiratorio, por lo que no son consideradas significativamente dañinas para la salud, sin embargo la exposición continua a altas concentraciones puede causar irritación de garganta y mucosas.

Las partículas  $PM_{10}$  y  $PM_{2.5}$  pueden contribuir a aumentar las enfermedades respiratorias como la bronquitis, enfermedades cardiovasculares y en general disminuir la esperanza de vida de adultos mayores y enfermos debido a que las partículas más pequeñas son más peligrosas para la salud de los seres humanos porque son capaces de alcanzar la zona inferior de los pulmones. Son retenidas en las vías respiratorias superiores, por lo que cerca de un tercio de ellas penetra hasta los pulmones. Su efecto depende de su composición química, pueden producir irritación de las vías respiratorias, agravar el asma y favorecer las enfermedades cardiovasculares. En el corto plazo la contaminación por  $PM_{10}$  puede causar el deterioro de la función respiratoria. En el largo plazo se asocia con el desarrollo de enfermedades crónicas, con el cáncer o con la muerte prematura.

Los grupos de la población con mayor susceptibilidad a los efectos de las partículas incluyen niños, ancianos, fumadores y personas con enfermedades respiratorias y cardiovasculares previas.



Figura 1.13: Molécula de Dióxido de Nitrógeno

### 1.3.3 Óxidos de Nitrógeno ( $NO_x$ )

Grupo de gases de color marrón-amarillento conformado por el nitrógeno y oxígeno que incluyen compuestos como óxido nítrico ( $NO$ ) y dióxido de nitrógeno ( $NO_2$ ) (figura 1.13). El término  $NO_x$  se refiere a la combinación de estas dos sustancias. Los procesos naturales y los realizados por el hombre producen óxidos de nitrógeno. Las fuentes más comunes de óxidos de nitrógeno en la naturaleza son los incendios forestales y de pastos, y la actividad volcánica. Las fuentes principales de emisión antropogénica son los escapes de los vehículos motorizados, las plantas eléctricas y la quema de combustibles fósiles.

#### Efectos en la Salud

El Dióxido de Nitrógeno ( $NO_2$ ) puede irritar los pulmones, causar bronquitis y pulmonía, así como reducción significativa de la resistencia respiratoria a las infecciones. Los efectos de exposición del  $NO_2$  a corto plazo no son claros, pero la exposición continua o frecuente a concentraciones mayores a las encontradas normalmente en el aire, puede causar un incremento en la incidencia de enfermedades respiratorias en los niños, agravamiento de afecciones en individuos asmáticos y con enfermedades respiratorias crónicas.

Actualmente se acepta que no hay evidencia científica confiable que sugiera la posibilidad de efectos crónicos atribuibles al  $NO_2$ , sin embargo éste puede ser fatal a concentraciones elevadas. Además, contribuye a la formación de lluvia ácida, que provoca graves daños a las plantas y materiales hechos por el hombre.

### 1.3.4 Ozono ( $O_3$ )

A temperatura y presión ambientales el ozono (figura 1.14), es un gas de olor acre e incoloro, que en grandes concentraciones puede volverse azulado. Es el que se encuentra en la zona más baja de la atmósfera y se forma a partir de ciertos precursores (óxidos de nitrógeno y compuestos orgánicos volátiles, como el formaldehído), contaminantes provenientes de la actividad humana.

#### Efectos en la Salud

Es un gas altamente oxidante que afecta a los tejidos vivos. Causa irritación ocular, de nariz y garganta. Provoca tos, y dificultad y dolor durante la respiración profunda. La exposición continua a este gas puede causar daño permanente a la membrana pulmonar. Atañe malestar general, debilidad, náusea y dolor de cabeza. La población de mayor riesgo a la contaminación



Figura 1.14: Molécula de Ozono



Figura 1.15: Molécula de Monóxido de Carbono

por ozono son los niños y adultos mayores, personas con enfermedades respiratorias y personas que realizan actividades al aire libre.

NOTA: Los daños por exposición al ozono dependen de la sensibilidad de cada individuo y del tipo de exposición.

El ozono tiene la singularidad de que es también beneficioso para los seres humanos y otros seres vivos. Es un componente necesario de la estratosfera (una capa elevada de la atmósfera que protege la capa inferior de la atmósfera), porque sirve para proteger a la tierra de la nociva radiación ultravioleta del sol. Sin embargo, cuando se encuentra en concentraciones altas en la capa inferior de la atmósfera, se le considera un contaminante.

### 1.3.5 Monóxido de Carbono (CO)

Es un gas incoloro, inodoro, inflamable y que en concentraciones altas puede ser letal (figura 1.15). La principal fuente antropogénica de monóxido de carbono es la quema incompleta de combustibles como la gasolina, keroseno, carbón, petróleo, tabaco o madera en ambientes de poco oxígeno. El monóxido de carbono es especialmente problemático en zonas urbanas con gran número de automóviles. El volumen del tránsito y el clima influyen sobre su concentración en el aire.

#### Efectos en la Salud

Dependen de la concentración y tiempo de exposición. El *CO* tiene la capacidad de unirse fuertemente a la hemoglobina, la proteína de los glóbulos rojos que contiene hierro y la cual se encarga de transportar el oxígeno a las células y tejidos a través de la sangre. Al combinarse





Figura 1.16: Instrumentos de Medición de Contaminantes Criterio



Figura 1.17: Instrumentos de Medición de Partículas  $PM_{10}$  y  $PM_{2.5}$

el  $CO$  con la hemoglobina, forma carboxihemoglobina ( $COHB$ ), lo cual indica una reducción significativa en la oxigenación de nuestro organismo (hipoxia), debido a que el  $CO$  tiene una afinidad de combinación 200 veces mayor que el oxígeno. La hipoxia causada por  $CO$  puede afectar el funcionamiento del corazón, del cerebro, de las plaquetas y del endotelio de los vasos sanguíneos. Su peligro es mayor en aquellas personas que padecen enfermedades cardiovasculares, angina de pecho o enfermedad vascular periférica.

Se le ha asociado con la disminución de la percepción visual, capacidad de trabajo, destreza manual y habilidad de aprendizaje. El peligro es más evidente en nonatos, neonatos, ancianos y en quienes sufren de enfermedades crónicas.

Observación: Su importancia radica en los daños que puede causar a la salud humana al permanecer expuestos por períodos prolongados a concentraciones elevadas de éste contaminante.

### 1.3.6 Normas de Calidad del Aire

Algunos de los instrumentos que miden los contaminantes criterio como  $SO_2$ ,  $NO_2$ ,  $CO$  y  $O_3$  se observan en la figura 1.16, y los equipos que miden las partículas  $PM_{10}$  y  $PM_{2.5}$  se muestran en la figura 1.17.

Para cada contaminante criterio se han establecido normas que establecen las concentraciones máximas permisibles de los contaminantes atmosféricos durante un período definido. Son los valores límite diseñados con un margen de protección ante los riesgos. La finalidad de las normas es proteger la salud humana y proteger el bienestar del ser humano y los ecosistemas. Las normas de calidad del aire para los contaminantes criterio se observan en la figura 1.18.

OZONO (O3)	Limite Máximo Permitido		Publicación
	Horario	Móvil de 8hrs.	
NOM- O2O-SSA1-1993	110 ppb no ser rebasado una vez al año	El valor quinto máximo debe ser menor o igual que 80 ppb	30 de Octubre de 2002
<b>BIOXIDO DE AZUFRE ( SO2 )</b>	<b>Promedio 24 hrs</b>	<b>Media anual</b>	<b>18 de agosto de 1994</b>
NOM- O22-SSA1-1993	130 ppb una vez al año	30 ppb media aritmética anual	
<b>MONOXIDO DE CARBONO (CO)</b>	<b>Móvil de 8hrs</b>		<b>12 de Agosto de 1994</b>
NOM- O21-SSA1-1993	No rebasar 11 ppm una vez al año		
<b>BIOXIDO DE NITROGENO (NO2 )</b>	<b>Bióxido de Nitrógeno</b>		<b>23 de Diciembre de 1994</b>
NOM- O21-SSA1-1993	No rebasar el valor máximo una vez al año de 210 ppb		
<b>PST, PM10, PM2.5</b> NOM- O25-SSA1-1993	<b>Promedio de 24hrs</b>	<b>Media Anual</b>	<b>26 de septiembre de 2005</b>
<b>PM10</b>	Percentil 98 menor o igual a 120 µgr/m³	Menor o igual que 50 µgr/m³	
<b>PM2.5</b>	<b>Promedio de 24hrs</b> Percentil 98 menor o igual a 65 µgr/m³	<b>Media Anual</b> Menor o igual que 15 µgr/m³	
<b>PST</b>	<b>Promedio de 24hrs</b> Percentil 98 menor o igual a 210 µgr/m³		

Figura 1.18: Normas de Calidad del Aire

### 1.4 Variables Meteorológicas

Del griego *μετεωρου*, *meteo*, que significa “alto en el cielo”; y *λογος*, *logos*, “conocimiento, tratado”. La meteorología es una ciencia interdisciplinaria que estudia el estado del tiempo, el medio atmosférico, los fenómenos allí producidos y las leyes que los rigen.

Hay que recordar que la tierra está constituida por tres partes fundamentales: una parte sólida llamada litósfera, recubierta en buena proporción por agua, y a esta capa se le llamada hidrósfera y, ambas, litósfera e hidrósfera, están envueltas por una tercera capa gaseosa llamada atmósfera. Estas capas se relacionan entre sí produciendo modificaciones profundas en sus características. La ciencia que estudia estas características, las propiedades y los movimientos de las tres capas fundamentales de la tierra es la geofísica. En ese sentido, la meteorología es una rama de la geofísica que tiene por objeto el estudio detallado de la envoltura gaseosa de la tierra y sus fenómenos. Mediante el estudio de los fenómenos que ocurren en la atmósfera la meteorología trata de definir el clima, predecir el tiempo y comprender la interacción de la atmósfera con otros subsistemas, etc.

La Meteorología aplicada tiene por objeto recolectar constantemente un máximo de datos sobre el estado de la atmósfera y analizarlos, interpretarlos y obtener deducciones prácticas, especialmente para prever el tiempo con la máxima antelación.

Como la atmósfera es una inmensa masa gaseosa sujeta a variaciones constantes que, la mayoría de las veces se producen en el ámbito regional, su estado, en un momento dado, sólo puede ser conocido si se dispone de una red suficientemente densa de puestos de observación o estaciones meteorológicas distribuidas por todas las regiones del globo, que a horas fijas efectúan las mismas mediciones (temperatura, presión, humedad, viento, precipitaciones, nubosidad, etc.)

TESIS TESIS TESIS TESIS TESIS

y transmiten los resultados a los centros encargados de utilizarlos.

Algunas de las variables meteorológicas de mayor interés para el Instituto de Ecología son la temperatura, velocidad y dirección del viento y la estabilidad atmosférica. A continuación se describen cada una de ellas.

#### 1.4.1 Presión Barométrica

También se conoce como presión atmosférica. Es la presión o el peso que ejerce la atmósfera en un punto determinado. La medición puede expresarse en varias unidades de medidas como hectopascales (*hPa*), milibares (*mbar*), milímetros o pulgadas de mercurio (*mmHg*, *inchHg*).

La presión barométrica a lo largo del año muestra el comportamiento meteorológico, por ejemplo una zona de baja presión típicamente está acompañada de nubosidad e incremento en el viento, mientras que una zona de alta presión típicamente ofrece cielo despejado.

La importancia del barómetro no radica en observar el valor numérico actual, más importante es la tendencia barométrica en las últimas horas o días. Esta tendencia es la que señala la presencia de los centros de alta o baja presión que se mueven sobre nosotros.

#### 1.4.2 Temperatura Ambiente

Es la temperatura que se puede medir con un termómetro y que se toma del medio ambiente actual, por lo que, si se toma de varios puntos en la tierra a un mismo tiempo puede variar.

De acuerdo a la clasificación propuesta y por los datos históricos acerca de la temperatura promedio en el estado de Guanajuato se sabe que en la mayor parte del estado y en las ciudades estudiadas, se tiene una temperatura templada.

#### 1.4.3 Humedad Relativa

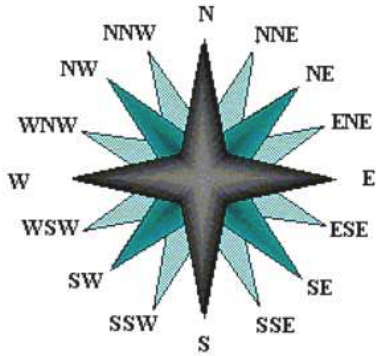
Es el cociente entre la humedad absoluta y la cantidad máxima de agua que admite el aire por unidad de volumen. Se mide como porcentaje y está normalizada de forma que la humedad relativa máxima posible es el 100%.

Una humedad relativa del 100% significa un ambiente en el que no cabe más agua. El cuerpo humano no puede transpirar y la sensación de calor puede llegar a ser asfixiante. Corresponde a un ambiente húmedo. Una humedad del 0% corresponde a un ambiente seco; se transpira con facilidad.

#### 1.4.4 Radiación Solar

Conjunto de radiaciones electromagnéticas emitidas por el sol y se distribuye desde el infrarrojo hasta el ultravioleta. No toda la radiación alcanza la superficie de la Tierra, pues las ondas ultravioletas más cortas, son absorbidas por los gases de la atmósfera fundamentalmente por el ozono.

La irradiancia es la magnitud utilizada para describir la potencia incidente por unidad de superficie de todo tipo de radiación electromagnética, es decir que mide la radiación solar que por unidad de tiempo y área, alcanza a la Tierra. En unidades del sistema internacional se mide en  $W/m^2$ . También se la utiliza para definir la constante solar, cantidad de energía solar que llega a la atmósfera superior de la Tierra por unidad de superficie.



NNE	Norte Noreste	22,50°
NE	Noreste	45,00°
ENE	Este Nordeste	67,50°
E	Este	90,00°
ESE	Este Sudeste	112,50°
SE	Sudeste	135,00°
SSE	Sur Sudeste	157,00°
S	Sur	180,00°
SSW	Sur Sudoeste	202,50°
SW	Sudoeste	225,00°
WSW	Oeste Sudeste	247,50°
W	Oeste	270,00°
WNW	Oeste Noroeste	292,50°
NW	Noroeste	315,00°
NNW	Norte Noroeste	337,50°
N	Norte	360,00°

Figura 1.19: Rosa de los Vientos

### 1.4.5 Dirección del Viento

#### Viento

Es el movimiento del aire. Los vientos globales se generan como consecuencia del desplazamiento del aire desde zonas de alta presión a zonas de baja presión.

#### Dirección del Viento

Punto del horizonte de donde viene o sopla el viento. Para distinguir uno de otro se les aplica el nombre de los principales rumbos de la brújula, según la conocida rosa de los vientos. Los cuatro puntos principales corresponden a los cardinales: Norte (*N*), Sur (*S*), Este (*E*) y Oeste (*W*). Se consideran hasta 32 entre estos y los intermedios, aunque los primordiales y más usados son los siguientes con su equivalencia en grados del cenit (ver figura 1.19).

### 1.4.6 Velocidad del Viento

Intensidad con la cual se desplaza el aire de un punto a otro. La velocidad del viento ayuda a la dispersión y al transporte de los contaminantes, por ello la importancia de estudiar su comportamiento.

Cuando hay aumento en la velocidad del viento, indica que la presión atmosférica está variando rápidamente y que habrá pronto cambios acelerados en el estado del tiempo. Por el contrario, una disminución en la velocidad del viento, o bien carencia de viento, indican una permanencia de las condiciones meteorológicas actuales.



Figura 1.20: Sensores Meteorológicos

#### 1.4.7 Sensores para medición de variables meteorológicas

Los sensores meteorológicos utilizados se listan a continuación. En la figura 1.20 se observan algunos de ellos.

Anemómetro: Dirección y velocidad del viento.

Barómetro : Presión barométrica.

Heliógrafo: Insolación del suelo.

Higrómetro: Humedad.

Piranómetro: Radiación solar.

Pluviómetro: Lluvia.

Termómetro: Temperatura.

### 1.5 Validación de la Información

El Instituto de Ecología del Estado de Guanajuato (*IEE*) ha implementado un Centro de Control en el que se concentran las bases de datos de la Red de Monitoreo Atmosférico del Estado de Guanajuato y los criterios de limpieza de datos que pasan por varios niveles de revisión para su validación final.

#### 1.5.1 Transferencia de datos de las estaciones al Centro de Control

En la Red de Monitoreo de la Calidad del Aire del Estado de Guanajuato existen dos tipos de transferencia de datos: por telemetría (vía módem) y por microondas (vía antena). Las redes de Irapuato, León, Silao y Celaya utilizan la primera y la de Salamanca puede transmitir mediante ambos sistemas. Los datos se almacenan en una base que está en el Centro de Control del Instituto de Ecología del Estado, en la ciudad de Guanajuato. Los datos de origen se encuentran en formato *.dat*, *.inx* o *.hly* y son sometidos al tratamiento de limpieza de datos para después ser mostrados en la página electrónica de la red de monitoreo de la calidad del aire de Guanajuato [CAG2009].

### 1.5.2 Validación de Datos

Después de haber sido generados por la red de monitoreo, los datos pasan por un primer nivel de revisión llamado limpieza de datos.

La primera etapa consiste en la limpieza de las bases de datos de aquellos valores etiquetados con banderas como no válidos, entre otras razones, por fallas en el sistema eléctrico, en los analizadores o de comunicación, así como por operaciones de calibración o mantenimiento de equipos; esta limpieza en su mayoría se realiza con software generado para este tipo de aplicación.

En la segunda etapa se eliminan los datos negativos o nulos debido a que la mayoría de los equipos de monitoreo cuentan con límites de detección inferiores (las concentraciones mínimas de un contaminante que el equipo de monitoreo puede registrar) que son mayores que cero. Cabe mencionar que se considera hasta un cierto valor de concentraciones negativas que son llevadas a cero, ya que estos valores no son sino el corrimiento en la curva de calibración del instrumento de medición que se deteriora con el paso del tiempo.

*Rango de Operación (Gross Limit).* Todos los equipos se ajustan a un rango de operación para detectar concentraciones entre umbrales preestablecidos.

La última etapa es la identificación del comportamiento fuera del patrón de los contaminantes, como son los datos extremos y las concentraciones constantes por un periodo de cuatro o más horas consecutivas. La gran parte de las veces que el equipo muestra alguna falla o eliminación de datos, se incorpora la abreviatura NA.

Para garantizar la calidad de la información generada en cada estación de monitoreo, la información será representativa sólo cuando se tenga el 75% de los datos, por tanto para que el promedio de 24 horas de un día sea válido necesitará contar con mínimo 18 horas; para que una hora sea válida necesitará contar con al menos 45 min. de los datos de la hora. El criterio aplica para cualquier tipo de dato.

La limpieza y validación de datos no siempre se puede realizar por programas o software estadísticos ya que no pueden cubrir en su totalidad el comportamiento de todos los contaminantes, por lo que la validación de estados estacionarios o con patrones erráticos muchas de las veces depende de los criterios y conocimientos que se tengan acerca del contaminante.

## Capítulo 2

# Metodología STATIS

Cuando se dispone de información sobre un conjunto de individuos y ésta información se puede estructurar en diferentes grupos de variables, es posible plantearse diversas preguntas al realizar un estudio simultáneo de estos grupos, entre otras: ¿tendrán más influencia unos grupos que otros?, ¿cómo equilibrar esta influencia?, ¿definen estos grupos diferentes estructuras sobre el mismo conjunto de individuos?, ¿se pueden comparar estas estructuras?, ..., preguntas que surgen de la propia estructuración de los datos. El método STATIS, desarrollado por el equipo de Y. Escoufier y L'Hermier des Plantes, aborda el estudio de información de esta naturaleza. Antes de introducirnos en la metodología STATIS, se describe el análisis de componentes principales, ya que es necesario dentro del método.

### 2.1 Análisis de Componentes Principales

El objetivo principal de un análisis de componentes principales es describir la variación en un conjunto de variables correlacionadas  $x_1, x_2, \dots, x_q$ , en términos de un nuevo conjunto de variables no correlacionadas  $y_1, y_2, \dots, y_q$ , cada una de las cuales es una combinación lineal de las variables originales. Las nuevas variables son derivadas en orden decreciente de *importancia*, en el sentido de que  $y_1$  representa la mayor cantidad de variación en los datos originales entre todas las combinaciones lineales de  $x_1, x_2, \dots, x_q$ . Así,  $y_2$  representa la mejor variación restante, a reserva de que no está correlacionada con  $y_1$ , y así sucesivamente, es decir, formando un sistema de coordenadas ortogonales. Las nuevas variables definidas por éste proceso,  $y_1, y_2, \dots, y_q$ , son los componentes principales.

La esperanza de un análisis de componentes principales es, que los primeros componentes representan una proporción importante de la variación en las variables originales, y puedan por tanto, ser usadas para proporcionar una dimensión inferior de éstas, para que sean útiles en una variedad razones. A continuación se describe como se obtienen los coeficientes que especifican las combinaciones lineales de las variables originales.

El primer componente principal de las observaciones,  $y_1$ , es la combinación lineal

$$y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1q}x_q,$$

cuya varianza muestral es la mayor entre todas las combinaciones lineales.

Dado que la varianza de  $y_1$  podría aumentar sin límite simplemente mediante el aumento de los coeficientes  $\mathbf{a}'_1 = (a_{11}, a_{12}, \dots, a_{1q})$ , escritos en forma de vector por conveniencia, una restricción debe ser colocada sobre estos coeficientes. Una limitación razonable es exigir que la suma de

los cuadrados de los coeficientes,  $\mathbf{a}'_1 \mathbf{a}_1$ , debería tomar el valor de 1, aunque son posibles otras limitaciones.

El segundo componente principal  $y_2 = \mathbf{a}'_2 \mathbf{x}$  con  $\mathbf{x} = (x_1, x_2, \dots, x_q)$ , es la combinación lineal con la mayor variación restante y sujeta a las siguientes 2 condiciones  $\mathbf{a}'_2 \mathbf{a}_2 = 1$  y  $\mathbf{a}'_2 \mathbf{a}_1 = 0$ . La segunda condición garantiza que  $y_1$  y  $y_2$  son incorrelacionadas. Similarmente, el  $j$ -ésimo componente principal es  $y_j = \mathbf{a}'_j \mathbf{x}$  es la combinación lineal que tiene la mayor variación restante, sujeta a las siguientes condiciones:  $\mathbf{a}'_j \mathbf{a}_j = 1$  y  $\mathbf{a}'_j \mathbf{a}_i = 0$  para  $(i < j)$ .

Para encontrar los coeficientes del primer componente principal, se necesita elegir los elementos del vector  $\mathbf{a}_1$ , que maximicen la variación de  $y_1$  respecto a la condición  $\mathbf{a}'_1 \mathbf{a}_1 = 1$ .

Para maximizar una función con varias variables sujetas a una o más condiciones, el método de *Multiplicadores de Lagrange* es usado. En este caso, esto conduce a la solución de que  $\mathbf{a}_1$  es el eigenvector de la matriz de varianzas-covarianzas muestral,  $\mathbf{S}$ , lo que corresponde a su mayor eigenvalor. Los otros componentes son derivados de la misma manera, con  $\mathbf{a}_j$  siendo el eigenvector de  $\mathbf{S}$  asociado con su  $j$ -ésimo mayor eigenvalor. Si los eigenvalores de  $\mathbf{S}$  son  $\lambda_1, \lambda_2, \dots, \lambda_q$ , entonces si  $\mathbf{a}'_j \mathbf{a}_j = 1$ , la variación del  $j$ -ésimo componente está dado por  $\lambda_j$ .

La variación total de los  $q$  componentes principales será igual a la variación total de las variables originales, tal que

$$\sum_{j=1}^q \lambda_j = s_1^2 + s_2^2 + \dots + s_q^2,$$

donde  $s_j^2$  es la varianza muestral de  $x_j$ . Lo anterior se puede escribir como

$$\sum_{j=1}^q \lambda_j = \text{traza}(\mathbf{S}).$$

En consecuencia, el  $j$ -ésimo componente principal representa la proporción  $P_j$  de la variación total de los datos originales, donde

$$P_j = \frac{\lambda_j}{\text{traza}(\mathbf{S})}.$$

Los primeros  $m$  componentes principales, donde  $m < q$ , representan la proporción acumulada  $P^{(m)}$ , donde

$$P^{(m)} = \frac{\sum_{j=1}^m \lambda_j}{\text{traza}(\mathbf{S})} = \sum_{j=1}^m P_j.$$

## 2.2 Metodología STATIS

El objetivo del método STATIS es estudiar la estructura común de las distintas tablas de datos, indicando la trayectoria de los individuos homogéneos, así como de los elementos heterogéneos, es decir, aquéllos que tienen un comportamiento diferente al resto.

La metodología STATIS (*Structuration des Tableaux A Trois Indices de la Statistique: Estructura Estadística de una Tabla de Tres Índices*), parte de un conjunto de tablas de datos (matrices cúbicas o triples, donde una o más de las entradas pueden ser el tiempo), obtenidas en situaciones



TESIS TESIS TESIS TESIS TESIS

diferentes. Se podría efectuar un análisis separado para cada tabla pero esto no permitiría tener una idea clara de la evolución de los datos, ya que los diferentes planos principales no se pueden superponer. Por esta razón es necesario tener un sistema de referencia común, llamado *compromiso*, en el cual se representen adecuadamente las diferentes estructuras aportadas por cada tabla.

Los objetivos de la metodología son 1) comparar y analizar las relaciones entre los diferentes conjuntos de datos, 2) combinarlos en una estructura común llamada *compromiso*, la cual se analiza por medio de componentes principales para revelar la estructura común entre las observaciones, y finalmente, 3) proyectar cada uno de los conjuntos de datos originales hacia el compromiso, para analizar comunalidades y diferencias [AHVD2007].

Para poder utilizar esta metodología, los datos de interés deben recolectarse en situaciones específicas. Cada una de éstas tiene una estrategia de análisis distinta. Las dos situaciones son:

1. Se recolectan medidas de los mismos individuos en diferentes ocasiones formando un arreglo para cada ocasión.
2. Las mismas variables se miden en  $p$  grupos distintos de individuos. En este caso, cada arreglo corresponde a un grupo distinto de individuos.

En la mayoría de los casos, los datos recolectados corresponden a las mismas variables medidas en los mismos individuos pero en situaciones distintas (por lo general en fechas diferentes). Para estos casos es posible analizar las variables o los individuos, de acuerdo a lo que se desea conocer. En caso de interesarse en los individuos, se utiliza la estrategia correspondiente a la situación número 1. Si se desea conocer a las variables, debe aplicarse la estrategia de la segunda situación.

El número y la naturaleza de las variables usadas para describir las observaciones puede variar de un conjunto de datos a otro, pero las observaciones deben ser las mismas en todos los conjuntos de datos.

Por ejemplo, los conjuntos de datos pueden ser medidas tomadas de las mismas observaciones (individuos u objetos) en ocasiones diferentes. En este caso, el primer conjunto de datos corresponde a los datos recolectados en el tiempo  $t = 1$ , el segundo de ellos corresponde a los datos recolectados en el tiempo  $t = 2$  y así sucesivamente. El objetivo de éste análisis es evaluar si la posición de las observaciones es estable con el tiempo.

Como otro ejemplo, los conjuntos de datos pueden ser medidas tomadas de las mismas observaciones por diferentes sujetos o grupos de sujetos. En este caso el primer conjunto de datos corresponde al primer sujeto, el segundo conjunto de datos corresponde al segundo sujeto y así sucesivamente. El objetivo de este análisis es evaluar si hay una concordancia entre los sujetos o grupos de sujetos.

La idea general detrás de la metodología STATIS es analizar la estructura de los conjuntos de datos individuales y derivar de esta estructura un conjunto óptimo de pesos llamado *compromiso*, para calcular una representación común de las observaciones. Los pesos utilizados para calcular el compromiso son seleccionados para hacerlos representativos de todos los conjuntos de datos como sea posible. El Análisis de Componentes Principales del compromiso da la posición de las observaciones en el espacio del compromiso. La posición de las observaciones para cada conjunto de datos puede representarse en el espacio del compromiso como puntos suplementarios. Finalmente, como un subproducto del cálculo de los pesos, los conjuntos de datos pueden representarse como puntos en un espacio multidimensional.

### 2.2.1 Notación

Como ya se dijo, la metodología STATIS estudia simultáneamente  $K$  tablas de datos  $Z_1, \dots, Z_K$  de individuos por variables continuas, donde los individuos permanecen fijos. En este caso cada tabla de datos  $Z_k$  es de tamaño  $n_k \times p$  y los tripletes  $(Z_k, M, D_k)$  se llaman estudios, donde  $M$  es una matriz  $p \times p$  definida positiva y  $D_k$  es la matriz diagonal de los pesos de los individuos. La matriz diagonal  $D_k$  es la matriz de pesos de los estudios (generalmente  $D_k = \text{diag}(\frac{1}{K})$ ).

La estrategia de la metodología STATIS consiste de tres etapas, y se presentan a continuación [CEGV1998].

### 2.2.2 Interestructura

Esta etapa tiene el objetivo de estudiar la diferenciación global entre las tablas de datos. Para eso se utiliza la distancia inducida por el producto escalar de Hilbert-Schmidt entre los operadores  $R_k M$ , donde se supone que  $R_k$  es la matriz de covarianzas o correlaciones de las columnas de  $Z_k$ , y  $M$  es una métrica euclídea en el espacio de individuos, con frecuencia igual a la identidad.

La matriz  $S$  de productos escalares de Hilbert-Schmidt es por definición

$$S_{ij} = \text{tr}(R_i M R_j M) = \sum_s (R_i M)(s)(R_j M)'(s),$$

donde  $(R_i M)(s)$  es la fila  $s$  de la matriz  $R_i M$ .

Se sabe que las filas de la matriz  $H = U \Delta_\lambda^{\frac{1}{2}}$  forman una imagen euclídea exacta de los operadores  $R_k M$ , donde  $U$  es la matriz cuyas columnas son los vectores propios de  $SD_k$ ,  $D_k$  ortonormados.  $\Delta_\lambda$  es la matriz diagonal de los valores propios correspondientes mayores que cero. Tomando los dos primeros vectores propios se obtiene una imagen euclidiana bidimensional de los operadores, la cual es óptima entre todas las de rango 2 y cuya representación (usando un referencial ortonormado) da los mapas de la interestructura.

### 2.2.3 El Compromiso

Consiste en calcular una configuración llamada compromiso que es representativa de las  $K$  configuraciones y cuyo papel es definir un escenario para la representación de las trayectorias de los individuos y las variables. La construcción de dichas configuraciones requiere la diagonalización del operador  $CM$  con

$$C = \sum_{k=1}^K \beta_k \frac{R_k}{\|R_k\|},$$

y  $\beta = (\beta_1, \dots, \beta_K)$  es un primer vector propio de  $SD_k$  tal que

$$\sum_{k=1}^K \beta_k = 1.$$

### 2.2.4 Intraestructura

Es la etapa en la cual se analiza la trayectoria de las variables asociadas a cada tabla de datos, a través de los  $K$  estudios. También se consideran las trayectorias de los individuos en los períodos

donde permanecen fijos. Lo anterior se completa con gráficas de las variables observadas en el período entero, llamadas supervariables.

Sea  $V$  la matriz cuyas columnas son los vectores propios  $M$  ortonormados de  $CM$ , entonces la matriz

$$\beta_k R_k (V \Delta_\lambda^{-\frac{1}{2}}),$$

contiene en la  $j$ -ésima fila las coordenadas de la variable suplementaria  $x_k^j$  donde  $\Delta_\lambda$  es la matriz diagonal de los valores propios de  $CM$ , mayores que cero. Por otra parte, el vector de coordenadas del individuo  $x_i(k)$  en los ejes del compromiso es  $x_i(k)MV$  donde  $x_i(k)$  es la  $i$ -ésima fila de la matriz estandarizada  $Z_k$ . A partir de esta información se construyen las gráficas bidimensionales de las trayectorias de las variables, la gráfica de las supervariables y cuando corresponde, las trayectorias de los individuos.

### 2.2.5 Estandarización de los Datos

Los datos que se analizan son  $K$  tripletes  $(Z_k, M, D_k)$  donde  $Z_k$  son las tablas de datos de tamaño  $n_k \times p$ , donde  $n_k$  es el número de individuos asociados a la  $k$ -ésima ocasión, y  $p$  es el número de variables.

La operación de estandarización de las matrices  $Z_k$  consiste en restar a cada entrada de cada columna, su media, y dividir entre su desviación estándar.

### 2.2.6 Matriz de Productos Escalares

Sea  $R_k$  la matriz de correlaciones de las columnas de  $Z_k$ , y  $S$  la matriz de productos escalares, entonces

$$S_{kl} = tr(R_k M R_l M) = \sum_j R_k M(j) (R_l M)'(j),$$

donde  $R_k M(j)$  es la  $j$ -ésima fila de la matriz  $R_k M$ .

En la interestructura se distinguirán dos tipos de representación de la interestructura: la 'no centrada normalizada' y la 'centrada no normalizada'.

La matriz de coordenadas de la interestructura no centrada normalizada es

$$H1 = U \Delta_\lambda^{-\frac{1}{2}}$$

donde  $U$  es la matriz cuyas columnas son los vectores propios de  $S n D k$ ,  $D_k$  ortonormados y  $\Delta_\lambda$  es la matriz diagonal de los valores propios correspondientes.

Similarmente para el caso de la interestructura 'centrada no normalizada', se calcula su matriz de coordenadas  $H2$ . Debe recordarse que la matriz  $S$  centrada es

$$S_c = (I - 11' D_k) S (I - D_k 11')$$

donde  $1 = (1 \dots 1)'_{K \times 1}$ .

Como se ha dicho, el compromiso es un promedio ponderado de las matrices de correlación  $R_k$  normalizadas:

$$C = \sum_{k=1}^K \beta_k \frac{R_k}{\|R_k\|},$$

donde  $\beta = (\beta_1 \dots \beta_K)$  es un primer vector propio de  $S_n D_k$  (matriz de productos escalares normalizados), tal que

$$\sum_k^K \beta_k = 1.$$

Sea  $V$  la matriz cuyas columnas son los vectores propios  $M$  ortonormados de  $CM$ , entonces la matriz

$$\beta_k R_k(V \Delta_\lambda^{-\frac{1}{2}}),$$

contiene en la  $j$ -ésima fila las coordenadas de la variable suplementaria  $x_k^j$ .



## Capítulo 3

# Aplicación

Como se ha indicado, la zona del Corredor Industrial del Estado de Guanajuato ha sido tipificada como crítica con respecto a la contaminación ambiental desde 1994. En el 2006 se realizó un informe de estado y tendencias de la contaminación del aire por parte de Instituto de Ecología del estado, donde se analizaron tanto contaminantes ambientales como variables meteorológicas. En dicho informe se mostró el comportamiento de cada contaminante y variable meteorológica durante un periodo específico. El propósito del estudio fue conocer el nivel de contaminantes criterio al que está expuesta la población e informar oportunamente sobre la calidad del aire, así como contar con elementos para estimar los efectos en la población y el medio ambiente.

Los municipios de Celaya, Salamanca, Irapuato, Silao y León fueron objeto de estudio en el informe del 2006. A partir de 10 estaciones fijas de la Red de Monitoreo de la Calidad del Aire del Estado de Guanajuato se recolectó información horaria de 2 tipos de variables, contaminantes ambientales y variables meteorológicas.

Este documento propone un análisis longitudinal del comportamiento de la información obtenida durante el año 2006 para 6 contaminantes ambientales y 5 variables meteorológicas, basado en la metodología STATIS, la cual tiene como objetivos encontrar una estructura común dentro de los arreglos de datos (intraestructura), y encontrar la relación existente entre arreglos (interestructura), para así conocer la evolución de la contaminación a lo largo del tiempo.

### 3.1 Imputación de Información

A partir del Informe de Estado y Tendencias en el 2006 [LAR2007], al Instituto de Ecología del estado de Guanajuato, le interesó conocer si el comportamiento de los contaminantes ambientales y de las variables meteorológicas fue el mismo durante los 12 meses del año. Para ello, utilizaremos la metodología STATIS que analiza datos cuantitativos a partir de arreglos de datos de 3 entradas, donde al menos una de esas entradas puede medirse en el tiempo. A su vez, está se basa en un análisis de componentes principales, donde su objetivo principal es encontrar una estructura común dentro de los arreglos de datos.

Cuando la red de monitoreo del Instituto de Ecología del estado de Guanajuato proporcionó los datos generados durante el 2006 al Centro de Investigación en Matemáticas, A. C. (CIMAT), Unidad Aguascalientes, lo hizo a través de archivos de texto (extensiones *.dat*) y hojas de cálculo de Microsoft Excel. Parte del proyecto *Estudio de la Calidad del Aire en el Estado de Guanajuato Mediante el Diseño de un Modelo Estadístico de Diagnóstico y Pronóstico* [PACR2006], fue depurar la información, presentando una base de datos de 79584 observaciones y 28 variables en formato

#	VARIABLE	TIPO DE DATO	VALORES TOMADOS	DESCRIPCIÓN
1	FECHA	Numérico	13355452800,-,13386902400	Código de 11 dígitos que indica la fecha en la que se realizó la lectura del dato.
2	TIEMPO	Numérico	0,3600,-,82800	Cantidad numérica que indica los segundos transcurridos a partir de la lectura del primer dato en 24 horas.
3	MES	Numérico	1,-,12	Cantidad numérica donde 1 es enero y 12 es diciembre.
4	DIA	Numérico	1,-,31	Cantidad numérica que indica el día del mes.
5	HORA	Numérico	0,-,23	Cantidad numérica que indica la hora del día.
6	COLUMNA	Numérico	1,-,24	Cantidad numérica que indica la medición del día.
7	O3	Numérico	-960.0,-, 499.6	Mediciones de Ozono.
8	PM10	Numérico	-960.00,-, 999.50	Mediciones de Material Particulado menor a 10µm.
9	CO	Numérico	-960.0,-, 34.5	Mediciones de Monóxido de Carbono.
10	SO2	Numérico	-960.0,-, 492.9	Mediciones de Dióxido de Azufre.
11	NO	Numérico	-960.0,-, 477.8	Mediciones de Óxido Nítrico.
12	NO2	Numérico	-960.0,-, 482.9	Mediciones de Dióxido de Nitrógeno.
13	NOX	Numérico	-960.0,-, 491.8	Mediciones de Óxidos de Nitrógeno.
14	CH4	Numérico	-999.00,-, 774.00	Mediciones de Metano.

Figura 3.1: Descripción de las variables de la Base de Datos (1)

SPSS.

Las variables incluidas se describen en las figuras 3.1 y 3.2.

Cabe resaltar que tanto en las variables de contaminación ambiental como en las variables meteorológicas existen muchos valores perdidos (*NA*), por lo que fue necesario analizar detalladamente cada una de las variables para incluirlas o no en el estudio (cuando un dato no está disponible se representa como *NA*, del inglés *Not Available*, independientemente del tipo de dato).

Debido a algunas bondades especiales del software R, como el manejo de grandes volúmenes de información, arreglos de datos y análisis gráficos (ver anexo A.1), se decidió programar en él. Después de la lectura de la base de datos en formato SPSS (extensión *.sav*), se eliminó la variable *FECHA*, debido a que era un código de 11 dígitos que indicaba la fecha en que se realizó la lectura de los datos (del 1 de enero de 2006 al 31 de diciembre de 2006). La variable numérica *TIEMPO*, que originalmente indicaba el número de segundos transcurridos desde la primera hora del día para cada medición, se transformó en una variable categórica con 8760 niveles, cada uno de los cuales formado por el mes, día y hora de la medición. A partir de lo anterior, la base de datos se transformó en un *data frame* (tabla de datos que integra variables de diferentes tipos) con 27 variables.

En las imágenes de descripción de los contaminantes ambientales y de las variables meteorológicas, la mayoría tienen valores negativos. En el proyecto mencionado anteriormente [PACR2006], se validó el límite de detección de los instrumentos de medición (*gross limit*).

Cuando una observación en las variables  $O_3$ ,  $SO_2$ ,  $NO$ ,  $NO_2$  y  $NO_X$  fue menor a  $-0.01$ , se asignó el valor  $-960$ , y si fue mayor a  $500$ , se asignó el valor  $-950$ . Si una observación en la variable  $CO$  fue menor a  $-0.01$ , se asignó el valor  $-960$ , y si fue mayor a  $50$ , se asignó el valor  $-950$ . Si una observación en la variable  $PM_{10}$  fue menor a  $-0.01$ , se asignó el valor  $-960$ , y si

#	VARIABLE	TIPO DE DATO	VALORES TOMADOS	DESCRIPCIÓN
15	NMHC	Numérico	-999.00,;, 153.50	Mediciones de Hidrocarburos no Metanos.
16	THC	Numérico	-999.00,;, 844.00	Mediciones de Hidrocarburos Totales.
17	WS	Numérico	-999.00,;, 64.50	Mediciones de la Velocidad del Viento.
18	WD	Numérico	-999.0,;, 825.8	Mediciones de la Dirección del Viento.
19	TEMP	Numérico	-999.00,;, 36.50	Mediciones de la Temperatura.
20	HR	Numérico	-999.00,;, 99.90	Mediciones de la Humedad Relativa.
21	PBAR	Numérico	-999.0,;, 850.0	Mediciones de la Presión Barométrica.
22	RADSOL	Numérico	-999.0,;, 999.8	Mediciones de la Radiación Solar.
23	TRIMESTRE	Numérico	1,;,4	Cantidades numéricas que indican el trimestre.
24	O3_2	Numérico	0.50,;,311.07	Promedios Móviles de Ozono debido a diferentes Normas de Calidad.
25	CO_2	Numérico	0.008750,;,15.59	Promedios Móviles de Monóxido de Carbono debido a diferentes Normas de Calidad.
26	SO2_2	Numérico	0.7083,;,225.2	Promedios Móviles de Dióxido de Azufre debido a diferentes Normas de Calidad.
27	ESTACION	Catagórico	'CEL POL',;, 'SAL NAT'	10 Niveles que representan cada estación de monitoreo.
28	CIUDAD	Catagórico	'CELAYA',;, 'LEON'	4 Niveles que representan 4 ciudades de Guanajuato.

Figura 3.2: Descripción de las variables de la Base de Datos (2)

fue mayor a 1000, se asignó el valor  $-950$ . Por último, si en alguna de las variables se tenían 4 valores consecutivos iguales, se asignó el valor  $-955$ .

En este estudio, aquellos valores menores a 0, se cambiaron por *NA*'s, pues es preferible tratarlos como valores perdidos y después imputarlos, que trabajar con mediciones negativas ilógicas en algunas de las variables. Por otro lado, de las 10 estaciones de monitoreo, la estación del DIF de Salamanca tiene solo 744 medidas horarias de las 8760 al año, por lo que se decidió eliminar a dicha estación del estudio.

Como la información del data frame se puede resumir por estación y ciudad, y además por trimestre, mes, día y hora, se realizaron arreglos de datos de 3 entradas para cada una de las 8 combinaciones posibles (trimestre - variables - estación, trimestre - variables - ciudad, mes - variables - estación, ..., hora - variables - ciudad), dichos arreglos arrojan el porcentaje de datos correctos. Después de revisar cada uno de los arreglos, y utilizando el análisis mensual por estación (en las figuras 3.3, 3.4 y 3.5, se muestran solo las estaciones de Irapuato), se decidió eliminar la variable PM10 que guarda la información del material particulado menor a  $10\mu m$ , debido a que en 5 de las 9 estaciones de monitoreo se tenía a lo más el 40% de información.

Observación. Las tablas y gráficas que no se muestran aquí, se encuentran al final de éste capítulo.

También se eliminaron las variables CH4, NMHC, THC, O3\_2, CO\_2 y SO2\_2. Las 3 primeras se eliminaron porque solo en la estación de la Cruz Roja de Salamanca se tenía información de los últimos 4 meses del año. Las últimas 3 variables no se incluyeron debido a que son promedios móviles de los contaminantes originales y no son útiles para nuestro trabajo. Otras variables que no se utilizaron son COLUMNA y WD, la primera de ellas solo indica la hora en la que se realizó la medición y la otra indica la dirección del viento, que para nuestro estudio, no es de interés.

Con ayuda de la variable TIEMPO mencionada anteriormente y un arreglo de 8760 horas, 12





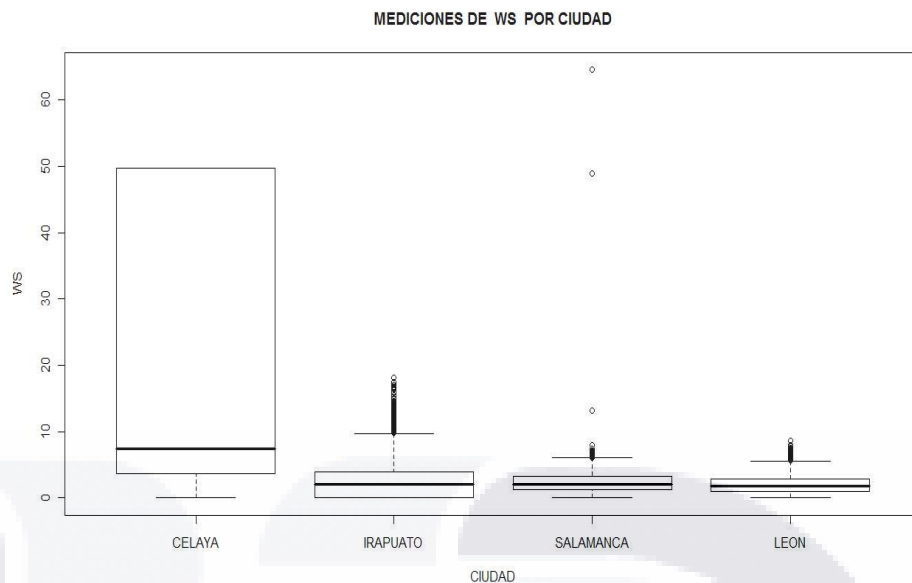


Figura 3.6: Caja y Brazos de la velocidad del viento en la ciudad de León

variables y 4 ciudades, se realizan gráficos de caja y brazos, histogramas y gráficos de acuerdo al orden de los contaminantes ambientales y de las variables meteorológicas. Por ejemplo, las gráficas de la velocidad del viento (WS) en León se muestran en las figuras 3.6, 3.7 y 3.8.

Recordemos que los óxidos de nitrógeno ( $NO_x$ ), se forman con la combinación de óxido nítrico ( $NO$ ) y dióxido de nitrógeno ( $NO_2$ ), por lo que analizando las 3 variables, se imputó de tal manera que si se tenían datos perdidos en alguna de ellas, por medio de las otras dos se pudiera obtener su valor.

Para imputar los datos faltantes en los contaminantes ambientales y en las variables meteorológicas se utilizó un arreglo de 12 meses, para las 12 variables y las 4 ciudades. En él, todo valor igual a cero o igual a  $NA$  se sustituyó por el valor del mes anterior.

Después de la eliminación de variables y de la imputación de información, se tiene un arreglo de datos de 3 entradas, donde las observaciones son las 4 ciudades de Guanajuato, las columnas son las 11 variables ambientales y meteorológicas, y las capas son los 12 meses del año. Además, el estadístico obtenido es el máximo de los datos, ya que nos interesan los valores mayores de cada variable. Ésta información se guarda en un archivo de datos llamado *arreglo\_cont.Rdata*, las figuras 3.9, 3.10 y 3.11, muestran la de los 3 primeros meses.

### 3.2 Análisis de Componentes Principales de la matriz de Correlaciones

Las variables numéricas de nuestro arreglo de datos tienen diferentes unidades de medida, por lo que deben ser estandarizadas. Los arreglos estandarizados de los primeros 2 meses se muestran en las figuras 3.12 y 3.13.

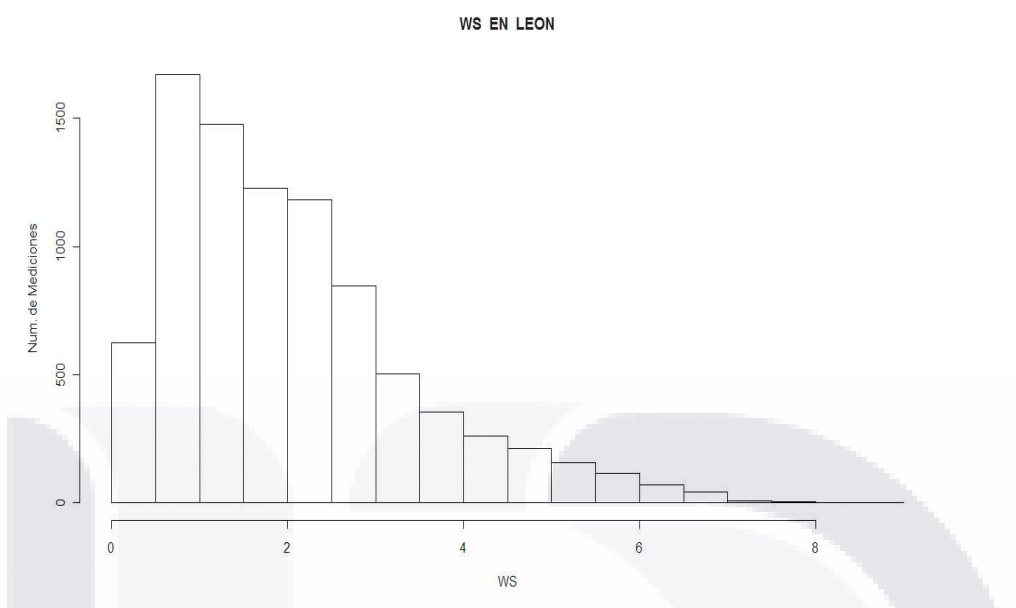


Figura 3.7: Histograma de la velocidad del viento en la ciudad de León

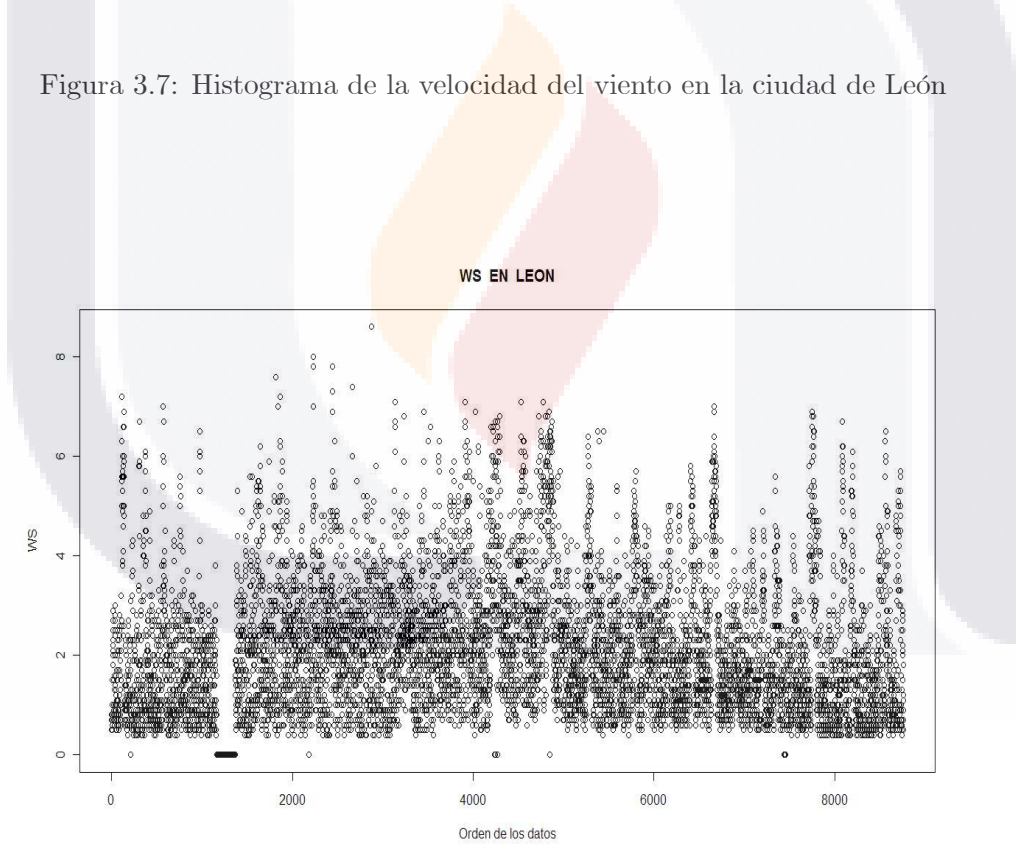


Figura 3.8: Gráfica de la velocidad del viento en la ciudad de León

Enero	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	499.6	7.6	79.0	222.1	201.4	223.0	14.2	27.2	99.5	680.5	886.3
Irapuato	254.4	8.8	399.5	271.4	62.1	280.0	13.8	26.8	99.9	681.4	873.7
Salamanca	88.0	8.8	446.0	226.5	48.6	259.8	6.4	26.7	93.2	834.7	768.9
León	283.7	3.2	101.9	338.2	394.0	459.5	7.2	28.8	86.6	673.0	967.3

Figura 3.9: Valores Máximos de las Variables en el mes de Enero

Febrero	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	271.9	13.2	70.0	188.9	124.5	231.5	49.7	31.0	82.3	680.3	947.7
Irapuato	96.7	6.7	218.9	167.8	54.8	210.3	13.9	29.5	79.7	681.3	950.0
Salamanca	73.0	4.8	349.2	172.2	36.6	190.1	6.7	30.3	74.8	832.7	768.9
León	302.3	3.2	106.6	152.9	248.0	375.6	6.5	32.0	71.0	672.4	997.1

Figura 3.10: Valores Máximos de las Variables en el mes de Febrero

Marzo	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	93.4	4.5	99.0	22.1	17.5	15.7	49.7	31.7	84.1	679.2	999.8
Irapuato	94.0	5.2	219.0	163.4	95.5	189.1	12.8	31.6	76.5	680.6	998.7
Salamanca	72.0	6.7	454.7	174.0	47.0	210.0	48.9	31.7	74.4	826.9	964.5
León	147.8	0.8	54.6	20.4	15.6	28.3	7.6	34.8	73.7	670.7	994.1

Figura 3.11: Valores Máximos de las Variables en el mes de Marzo

Enero	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	1.49	0.22	-1.06	-0.91	0.18	-0.91	1.05	-0.21	0.87	-0.54	0.17
Irapuato	-0.18	0.74	0.85	0.15	-0.82	-0.28	0.94	-0.68	0.94	-0.53	0.00
Salamanca	-1.32	0.74	1.13	-0.81	-0.92	-0.50	-1.11	-0.80	-0.29	1.73	-1.49
León	0.02	-1.69	-0.93	1.58	1.56	1.69	-0.89	1.69	-1.51	-0.65	1.32

Figura 3.12: Datos Estandarizados del mes de Enero

Febrero	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	0.84	1.64	-1.07	1.44	0.10	-0.28	1.71	0.33	1.23	-0.54	0.36
Irapuato	-0.87	-0.07	0.30	-0.21	-0.74	-0.57	-0.30	-1.31	0.63	-0.53	0.39
Salamanca	-1.11	-0.57	1.50	0.14	-0.96	-0.85	-0.70	-0.44	-0.49	1.73	-1.69
León	1.14	-0.99	-0.73	-1.37	1.59	1.70	-0.71	1.41	-1.36	-0.66	0.93

Figura 3.13: Datos Estandarizados del mes de Febrero

Enero	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	-0.20	-0.81	-0.07	0.41	-0.15	0.71	0.22	0.37	-0.75	0.56
CO	-0.20	1.00	0.70	-0.84	-0.97	-0.91	0.40	-1.00	0.80	0.47	-0.83
SO <sub>2</sub>	-0.81	0.70	1.00	-0.32	-0.85	-0.35	-0.19	-0.71	0.24	0.67	-0.78
NO	-0.07	-0.84	-0.32	1.00	0.73	0.96	-0.33	0.85	-0.70	-0.51	0.79
NO <sub>2</sub>	0.41	-0.97	-0.85	0.73	1.00	0.79	-0.24	0.98	-0.68	-0.57	0.87
NO <sub>x</sub>	-0.15	-0.91	-0.35	0.96	0.79	1.00	-0.54	0.91	-0.86	-0.33	0.71
WS	0.71	0.40	-0.19	-0.33	-0.24	-0.54	1.00	-0.37	0.86	-0.60	0.16
TEMP	0.22	-1.00	-0.71	0.85	0.98	0.91	-0.37	1.00	-0.78	-0.50	0.85
HR	0.37	0.80	0.24	-0.70	-0.68	-0.86	0.86	-0.78	1.00	-0.12	-0.35
PBAR	-0.75	0.47	0.67	-0.51	-0.57	-0.33	-0.60	-0.50	-0.12	1.00	-0.88
RADSOL	0.56	-0.83	-0.78	0.79	0.87	0.71	0.16	0.85	-0.35	-0.88	1.00

Figura 3.14: Matriz de Correlaciones de las Variables en Enero

Febrero	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.24	-0.91	-0.08	0.90	0.78	0.42	0.88	-0.13	-0.67	0.72
CO	0.24	1.00	-0.47	0.91	-0.20	-0.40	0.98	-0.13	0.90	-0.30	0.15
SO <sub>2</sub>	-0.91	-0.47	1.00	-0.10	-0.73	-0.60	-0.61	-0.61	-0.21	0.87	-0.87
NO	-0.08	0.91	-0.10	1.00	-0.50	-0.68	0.85	-0.31	0.86	0.12	-0.27
NO <sub>2</sub>	0.90	-0.20	-0.73	-0.50	1.00	0.98	-0.02	0.92	-0.51	-0.59	0.71
NO <sub>x</sub>	0.78	-0.40	-0.60	-0.68	0.98	1.00	-0.23	0.86	-0.65	-0.53	0.67
WS	0.42	0.98	-0.61	0.85	-0.02	-0.23	1.00	0.06	0.81	-0.38	0.26
TEMP	0.88	-0.13	-0.61	-0.31	0.92	0.86	0.06	1.00	-0.53	-0.29	0.42
HR	-0.13	0.90	-0.21	0.86	-0.51	-0.65	0.81	-0.53	1.00	-0.24	0.06
PBAR	-0.67	-0.30	0.87	0.12	-0.59	-0.53	-0.38	-0.29	-0.24	1.00	-0.98
RADSOL	0.72	0.15	-0.87	-0.27	0.71	0.67	0.26	0.42	0.06	-0.98	1.00

Figura 3.15: Matriz de Correlaciones de las Variables en Febrero

Para conocer las relaciones existentes entre las variables, se obtienen las matrices de correlaciones para cada mes. Las figuras 3.14 y 3.15 muestran las correlaciones de los primeros 2 meses.

Se desea analizar variables con distintas varianzas, de tal forma que se puedan agrupar cumpliendo 2 características: reducir la información sin perder de vista la dispersión de los datos y las relaciones entre ellos; y establecer un orden entre las observaciones, en este caso, entre las ciudades de Guanajuato. Un método que cumple lo anterior es el Análisis de Componentes Principales.

Después de aplicar el Análisis de Componentes Principales a las matrices de correlaciones de los arreglos por mes, se obtuvo una tabla resumen (ver figura 3.16), que muestra los valores de los primeros 3 eigenvalores, el porcentaje de variación explicada por cada uno y el porcentaje acumulado por los 3 eigenvalores, para cada mes del año. Se observa que en todos los meses los 3 primeros componentes explican el 100% de la información, y solo los dos primeros eigenvalores explican casi toda la variación en los datos.

La información de los eigenvalores también puede representarse gráficamente mediante un diagrama de ladera, el cual muestra en el eje horizontal el número del eigenvalor, y en el eje vertical su valor correspondiente. Las gráficas de ladera de los meses del año se muestran en las figuras 3.17, 3.18 y 3.19.

	Componente	Eigenvalor	Porcentaje	% Acumulado
Enero	1	6.84	62.20	62.20
	2	3.35	30.40	92.60
	3	0.81	7.40	100.00
Febrero	1	5.67	51.50	51.50
	2	4.27	38.80	90.30
	3	1.06	9.70	100.00
Marzo	1	6.64	60.30	60.30
	2	2.52	22.90	83.20
	3	1.84	16.80	100.00
Abril	1	6.17	56.10	56.10
	2	3.29	29.90	86.00
	3	1.54	14.00	100.00
Mayo	1	6.29	57.20	57.20
	2	3.33	30.30	87.50
	3	1.38	12.50	100.00
Junio	1	5.70	51.80	51.80
	2	3.52	32.00	83.80
	3	1.78	16.20	100.00
Julio	1	5.65	51.30	51.30
	2	3.58	32.60	83.90
	3	1.77	16.10	100.00
Agosto	1	5.30	48.10	48.10
	2	3.24	29.40	77.50
	3	2.47	22.40	99.90
Septiembre	1	6.06	55.10	55.10
	2	3.35	30.50	85.60
	3	1.58	14.40	100.00
Octubre	1	5.02	45.60	45.60
	2	3.62	32.90	78.50
	3	2.37	21.50	100.00
Noviembre	1	5.62	51.10	51.10
	2	3.68	33.40	84.50
	3	1.70	15.40	99.90
Diciembre	1	6.39	58.10	58.10
	2	3.07	27.90	86.00
	3	1.54	14.00	100.00

Figura 3.16: Eigenvalores de R y su Pocentaje de Variación para cada mes del año

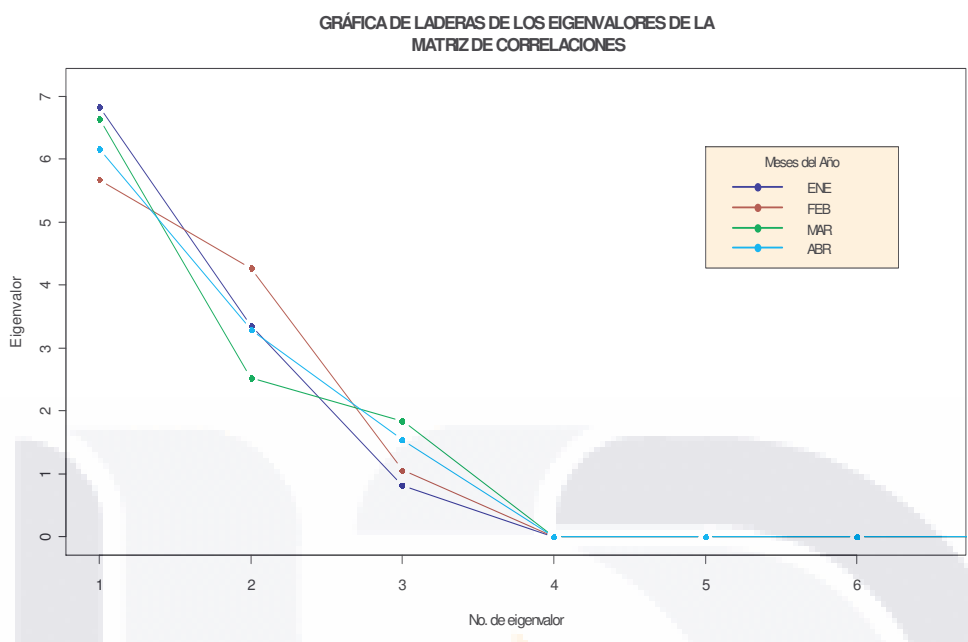


Figura 3.17: Gráficas de Laderas de Enero a Abril de 2006

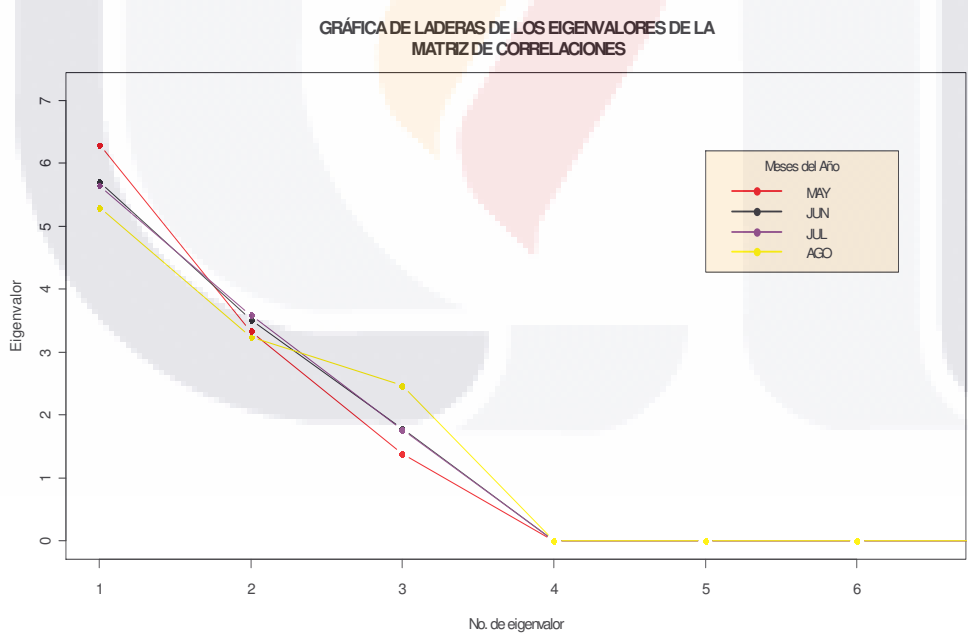


Figura 3.18: Gráficas de Laderas de Mayo a Agosto de 2006

GRÁFICA DE LADERAS DE LOS EIGENVALORES DE LA MATRIZ DE CORRELACIONES

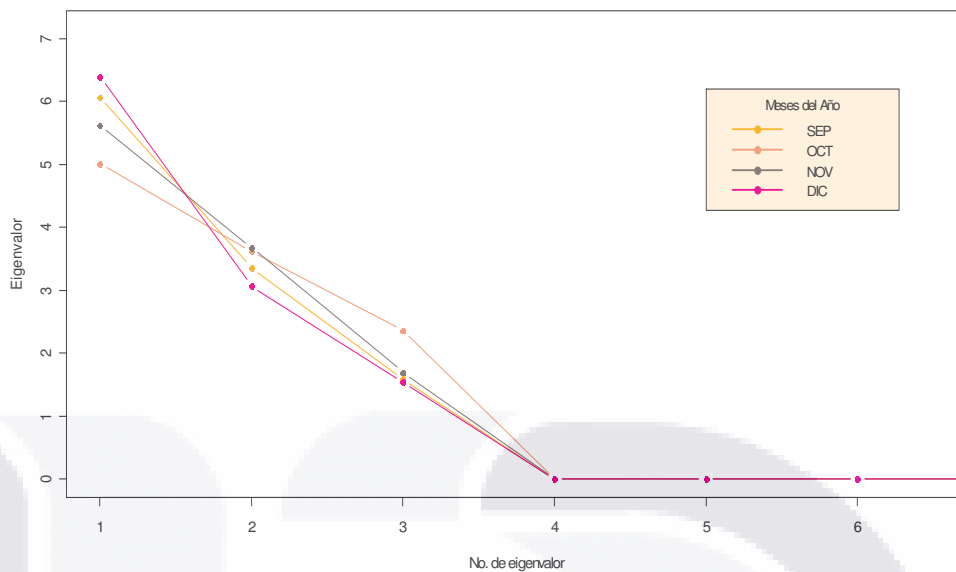


Figura 3.19: Gráficas de Laderas de Septiembre a Diciembre de 2006

El Análisis de Componentes Principales transformó el espacio de los vectores de la matriz de Correlaciones  $Z$ , en uno nuevo, es decir, se obtuvieron  $y_k \quad \forall k = 1, \dots, m$  con  $m \leq 11$  combinaciones lineales de las variables estandarizadas.

La tabla 3.1 representa las nuevas variables transformadas correspondientes al primer componente principal de cada mes. La tabla 3.2 muestra las variables transformadas correspondientes al segundo componente principal de cada mes.

Las nuevas variables  $y_k$  cumplen las siguientes características: no están correlacionadas,  $y_1$  tiene la mayor varianza,  $y_2$  tiene la segunda mayor varianza y así sucesivamente, es decir,  $Var(y_1) \geq Var(y_2) \geq \dots \geq Var(y_m)$ , y cada vector de puntajes está normalizado.

Para nuestro estudio, el primer componente estandarizado es el importante, debido a que éste es una medida relativa que describe la intensidad de las variables en un momento dado y además es el que contiene la mayor variación de la información.

Los puntajes del primer componente principal para cada mes se muestran en la figura 3.20, y los puntajes del segundo componente principal para cada mes se muestran en la figura 3.21. Los resultados arrojados pueden analizarse para cada mes, comparando en cada mes la posición de las ciudades de Guanajuato.

Las gráficas correspondientes a los puntajes de los primeros 3 meses del año se muestran en las figuras 3.22, 3.23 y 3.24. Además, en las figuras 3.25, 3.26, 3.27 y 3.28 se muestran los mismos puntajes, pero por ciudad, de acuerdo al transcurso de los meses.

Las gráficas de las correlaciones entre variables para los primeros 3 meses se muestran en las figuras 3.29, 3.30 y 3.31. Además, las gráficas de las correlaciones de las primeras 3 variables por mes se ven en las figuras 3.32, 3.33 y 3.34.

En cada uno de los meses se calculó el primer componente principal como combinación de 11 variables involucradas. La ponderación que se le da a cada una de éstas es distinta para cada mes,

	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>X</sub>	WS	TEMP	HR	PBAR	RADSOL
ENE	0.11	-0.50	-0.30	-0.24	0.14	0.44	0.54	0.005	-0.23	0.04	0.18
FEB	0.39	-0.13	-0.26	0.39	0.46	0.19	0.23	0.49	-0.18	-0.18	0.10
MAR	0.35	-0.26	-0.08	0.10	0.49	-0.24	0.25	0.20	0.57	-0.23	-0.14
ABR	0.37	-0.22	0.03	0.004	0.004	0.10	0.23	-0.39	0.63	0.20	-0.40
MAY	0.35	-0.26	0.04	-0.30	0.05	0.50	-0.31	0.02	0.54	0.11	0.25
JUN	0.20	-0.19	0.61	0.19	-0.38	0.46	-0.08	-0.19	0.22	0.28	-0.03
JUL	-0.27	-0.40	-0.002	-0.26	0.07	0.45	0.58	0.04	0.09	0.35	0.14
AGO	-0.35	-0.31	0.10	-0.67	-0.30	0.22	-0.04	-0.01	0.09	-0.27	0.31
SEP	0.20	0.44	-0.25	0.19	0.14	-0.45	0.11	0.07	0.58	-0.23	0.17
OCT	0.29	0.35	0.24	-0.10	-0.38	-0.43	-0.44	-0.28	0.16	0.13	-0.29
NOV	0.42	-0.02	-0.09	-0.22	0.14	0.30	0.03	-0.32	0.14	0.62	0.39
DIC	-0.14	-0.37	0.54	-0.34	-0.02	-0.44	0.25	-0.04	0.02	0.40	0.08

Tabla 3.1: Primer Componente Principal de cada mes

	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>X</sub>	WS	TEMP	HR	PBAR	RADSOL
ENE	-0.38	-0.07	0.11	0.06	0.14	0.40	-0.04	-0.23	0.37	-0.68	-0.03
FEB	-0.03	-0.48	-0.13	0.43	-0.07	0.08	-0.16	-0.06	0.26	0.32	-0.60
MAR	-0.37	0.19	0.10	0.14	0.50	0.12	-0.34	0.22	0.19	0.14	0.56
ABR	0.39	-0.11	0.003	-0.45	-0.10	0.24	0.23	0.50	-0.04	0.43	0.27
MAY	0.26	0.41	-0.13	0.05	-0.40	-0.33	-0.54	-0.31	0.17	-0.26	0.001
JUN	0.40	-0.01	-0.20	-0.32	-0.20	-0.12	-0.07	0.46	-0.08	0.65	0.02
JUL	0.38	-0.21	0.08	0.28	0.35	-0.30	0.18	0.67	0.10	0.13	0.09
AGO	-0.32	-0.38	0.06	0.20	0.02	-0.29	0.02	0.01	-0.76	-0.21	-0.06
SEP	0.37	-0.22	-0.02	0.09	0.14	-0.52	-0.08	-0.27	-0.49	-0.40	-0.19
OCT	-0.27	0.41	0.12	-0.29	-0.04	-0.16	0.32	0.49	0.20	-0.39	-0.31
NOV	0.31	-0.22	0.41	-0.19	0.22	-0.38	0.49	0.26	-0.05	0.24	-0.31
DIC	-0.29	0.25	0.42	0.09	0.15	0.15	0.44	0.15	0.43	-0.48	0.04

Tabla 3.2: Segundo Componente Principal de cada mes

PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL												
CIUDAD	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Celaya	-0.15	0.28	0.43	-0.84	1.53	-0.96	-1.03	0.76	-0.69	1.69	-1.08	0.41
Irapuato	-0.53	-0.50	-0.35	1.71	0.25	0.12	-0.38	-1.07	-0.90	-0.92	-0.50	-1.31
Salamanca	-0.98	-1.24	-1.40	-0.40	-0.73	-0.75	-0.25	1.21	1.65	-0.30	1.60	-0.47
León	1.66	1.46	1.32	-0.46	-1.04	1.58	1.65	-0.90	-0.06	-0.47	-0.02	1.38

Figura 3.20: Puntajes del Primer Componente Principal por mes



CIUDAD	PUNTAJES DEL SEGUNDO COMPONENTE PRINCIPAL											
	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Celaya	-1.37	-1.59	1.68	-0.51	-0.16	-0.94	-0.78	-0.24	1.42	-0.35	0.99	-0.99
Irapuato	-0.44	-0.12	-0.50	-0.13	-0.04	-0.83	-0.58	-1.32	-1.39	-1.06	-0.02	-0.72
Salamanca	1.30	0.78	-0.24	1.65	1.50	1.54	1.71	0.08	-0.17	1.64	0.64	1.57
León	0.51	0.92	-0.94	-1.01	-1.31	0.23	-0.36	1.48	0.15	-0.23	-1.62	0.14

Figura 3.21: Puntajes del Segundo Componente Principal por mes

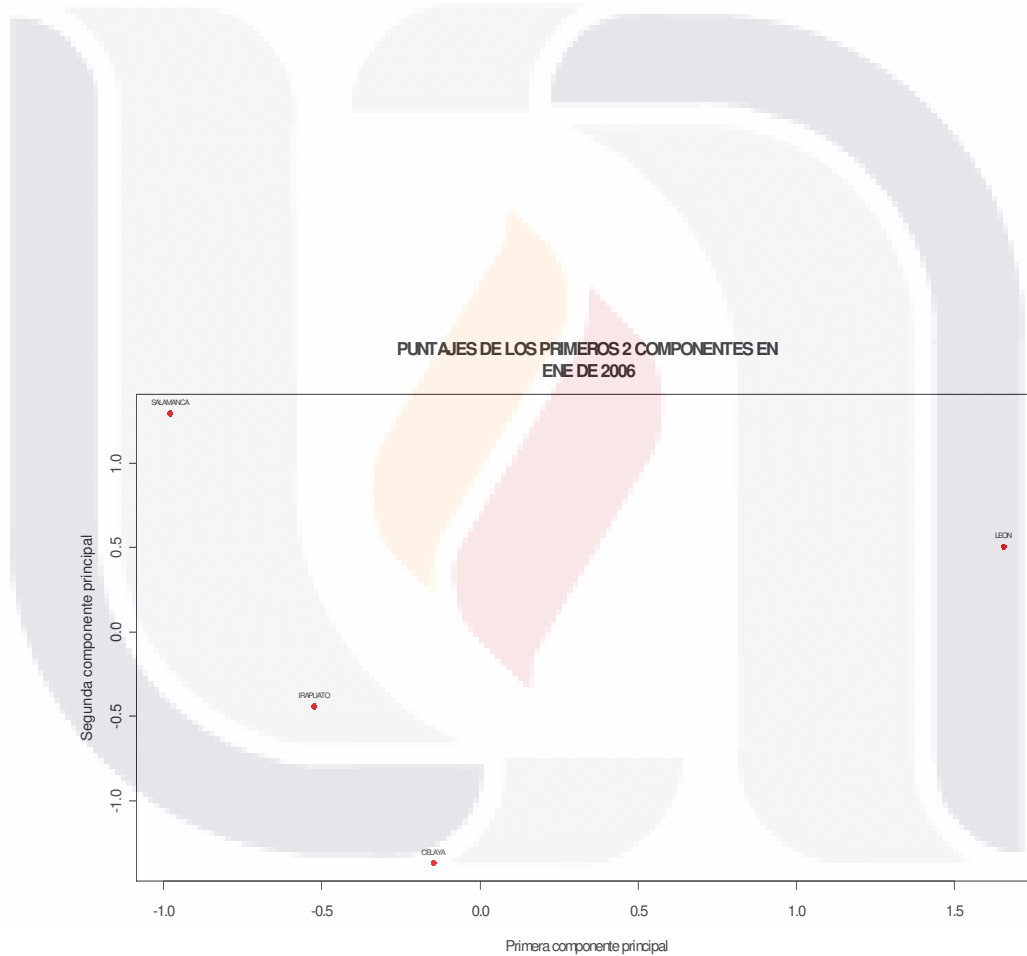


Figura 3.22: Puntajes de las Ciudades de Guanajuato en Enero

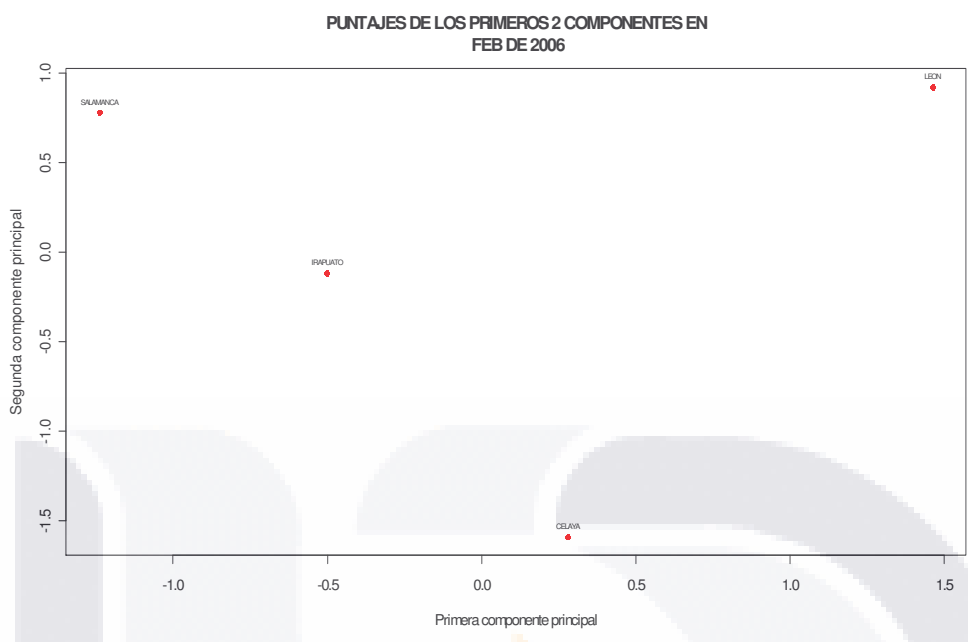


Figura 3.23: Puntajes de las Ciudades de Guanajuato en Febrero

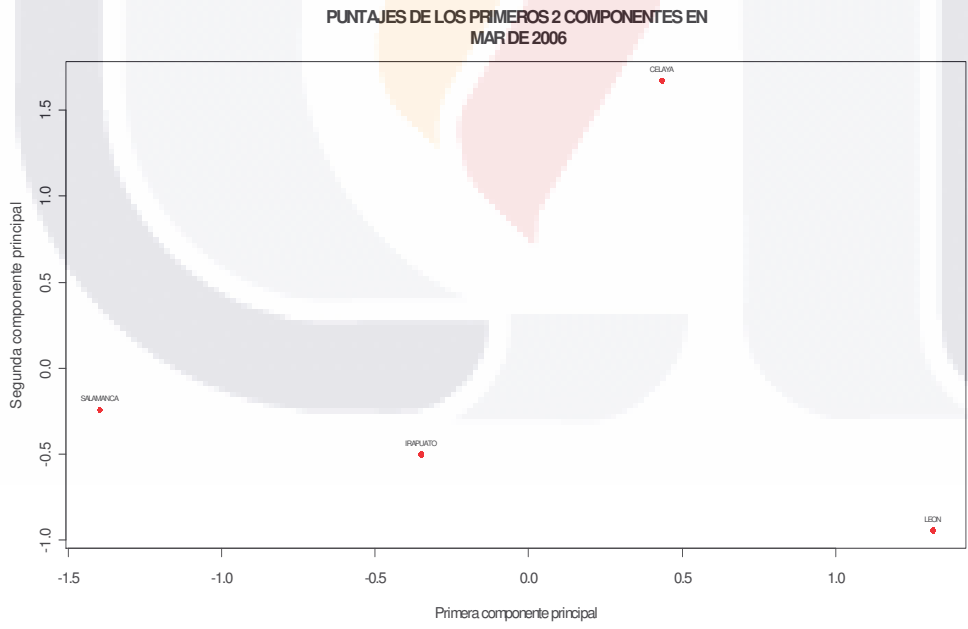


Figura 3.24: Puntajes de las Ciudades de Guanajuato en Marzo

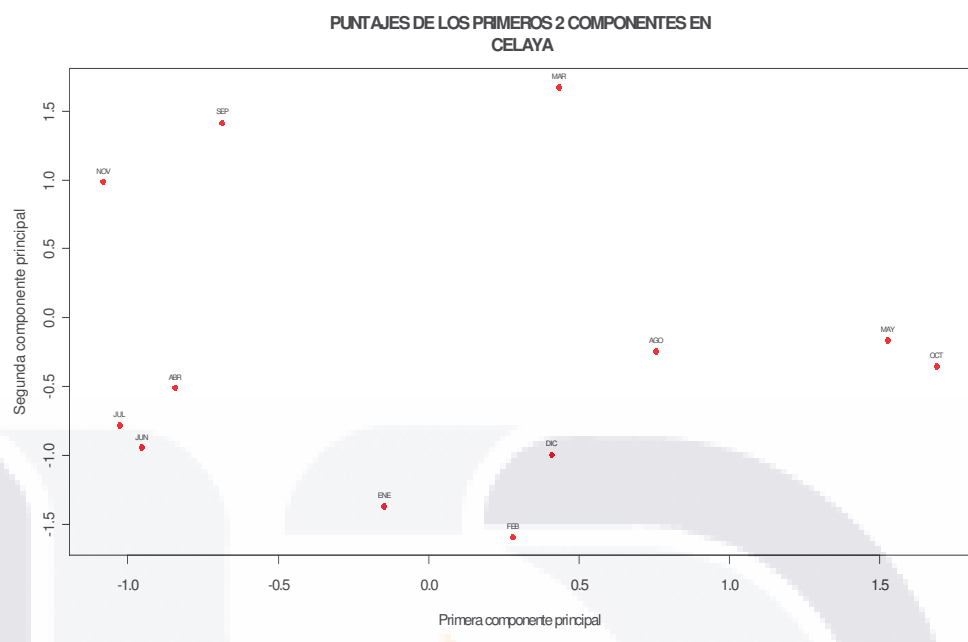


Figura 3.25: Puntajes Mensuales en Celaya

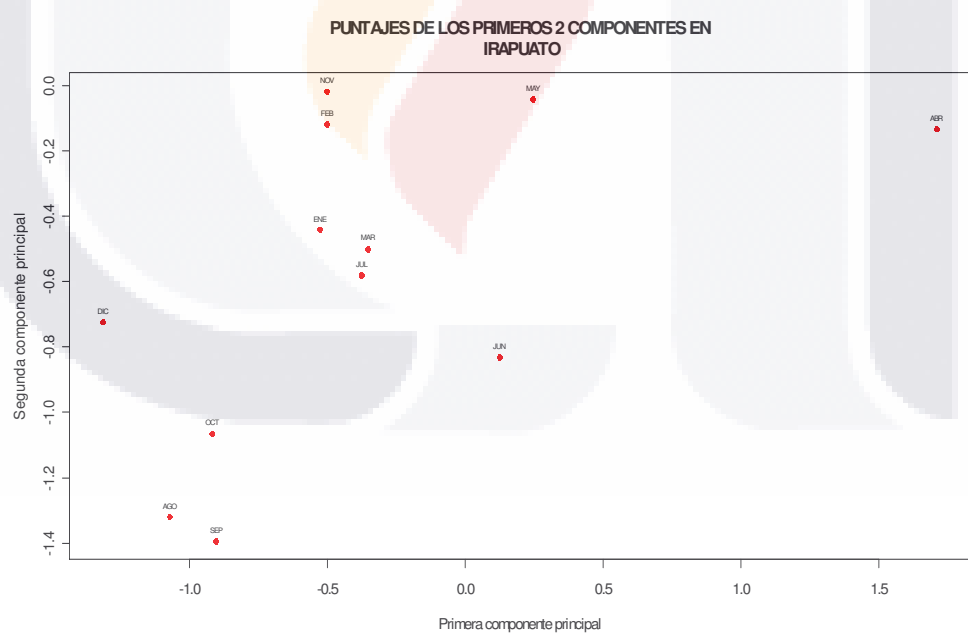


Figura 3.26: Puntajes Mensuales en Irapuato

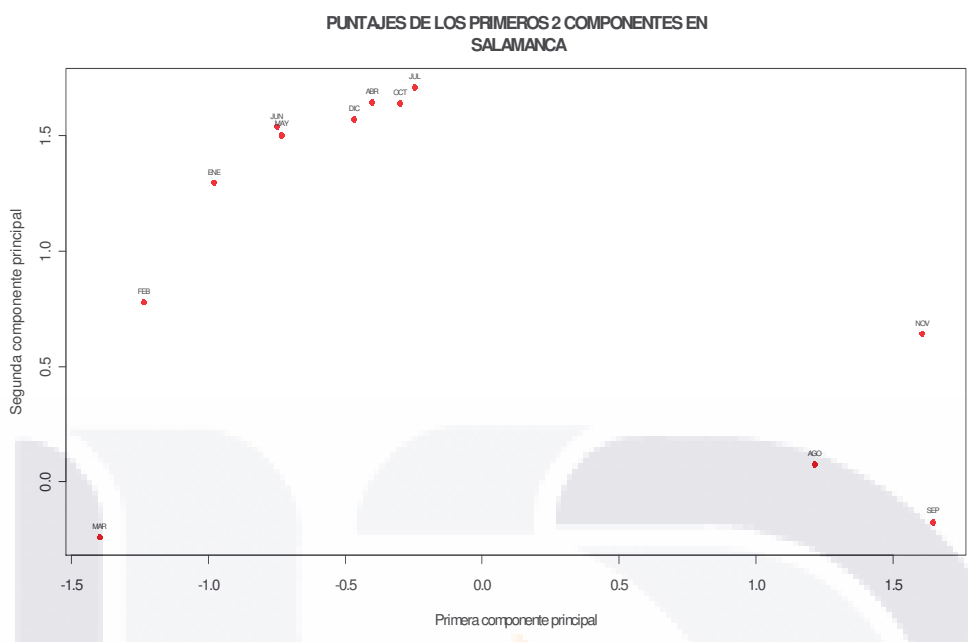


Figura 3.27: Puntajes Mensuales en Salamanca

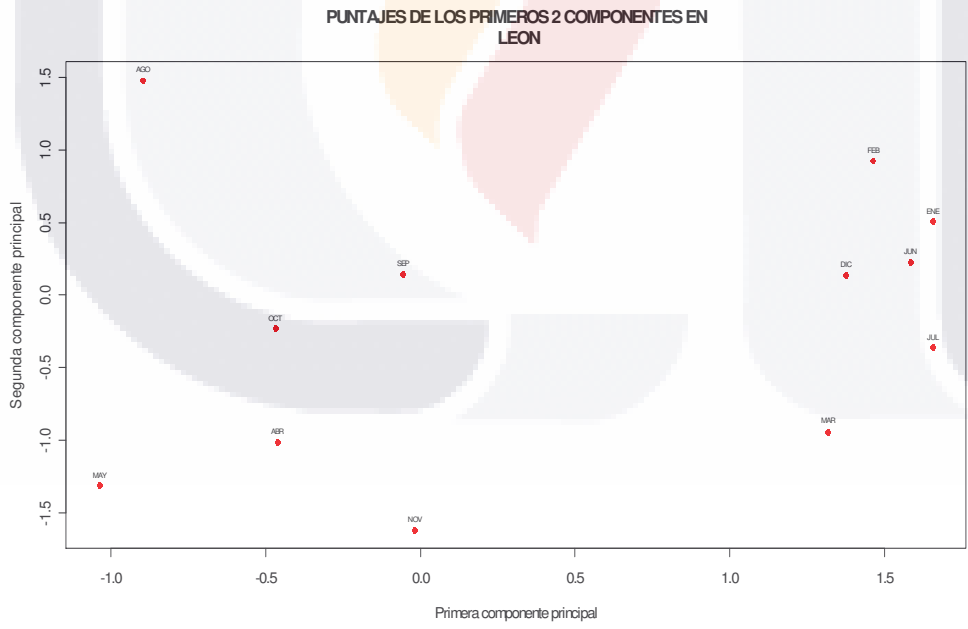


Figura 3.28: Puntajes Mensuales en León

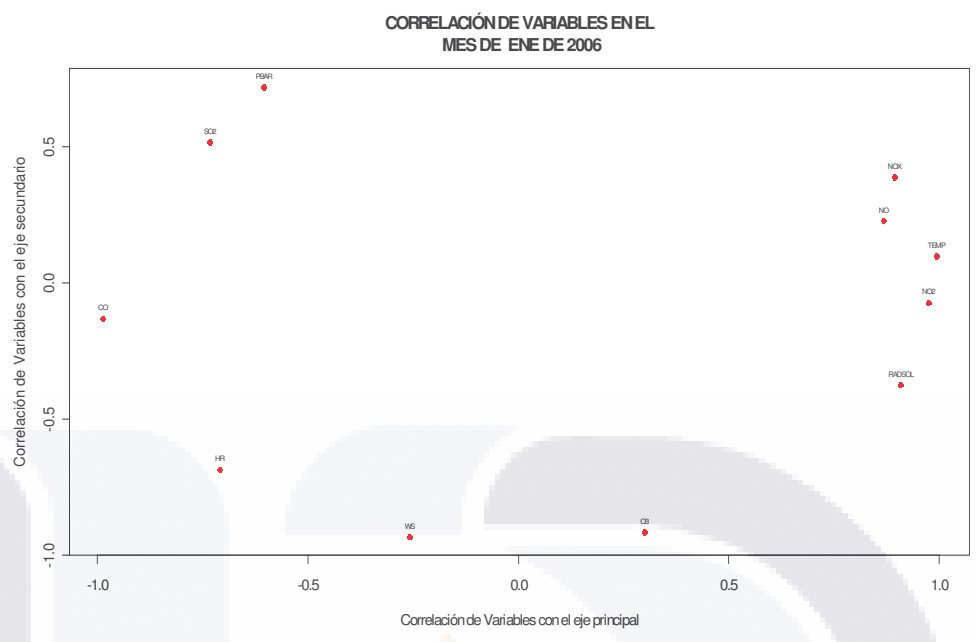


Figura 3.29: Correlaciones entre Variables en Enero

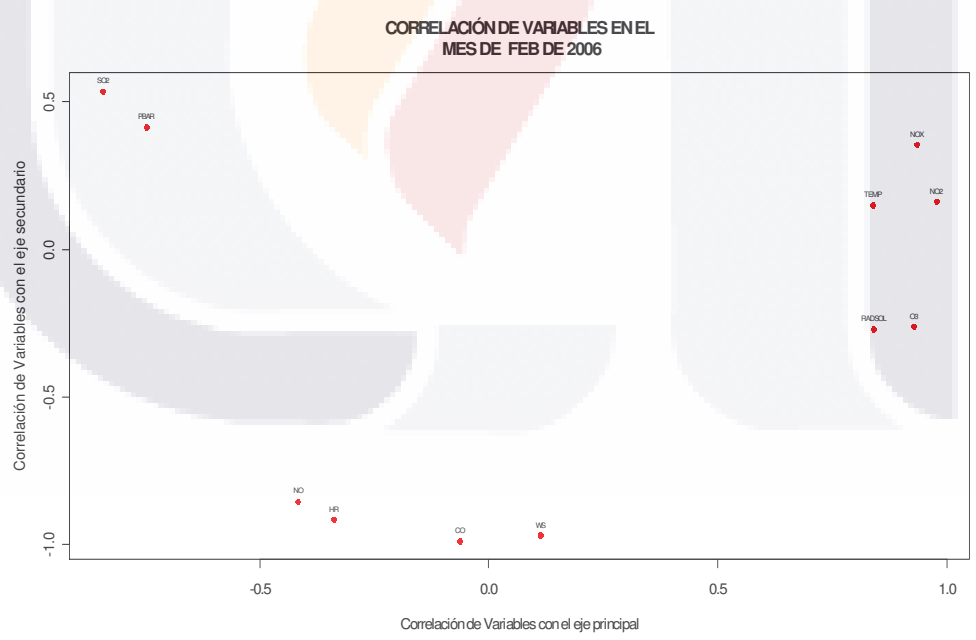


Figura 3.30: Correlaciones entre Variables en Febrero

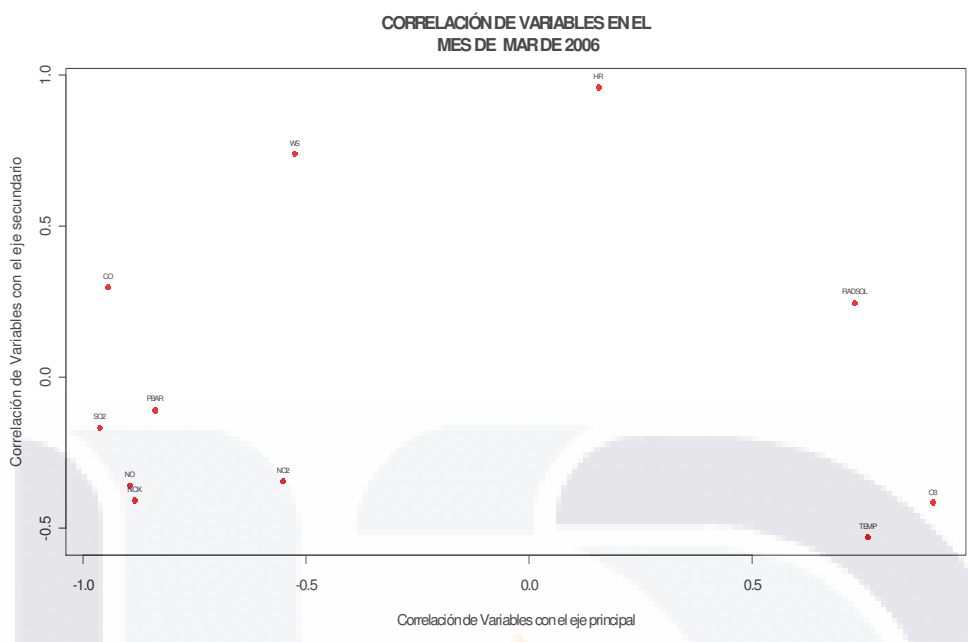


Figura 3.31: Correlaciones entre Variables en Marzo

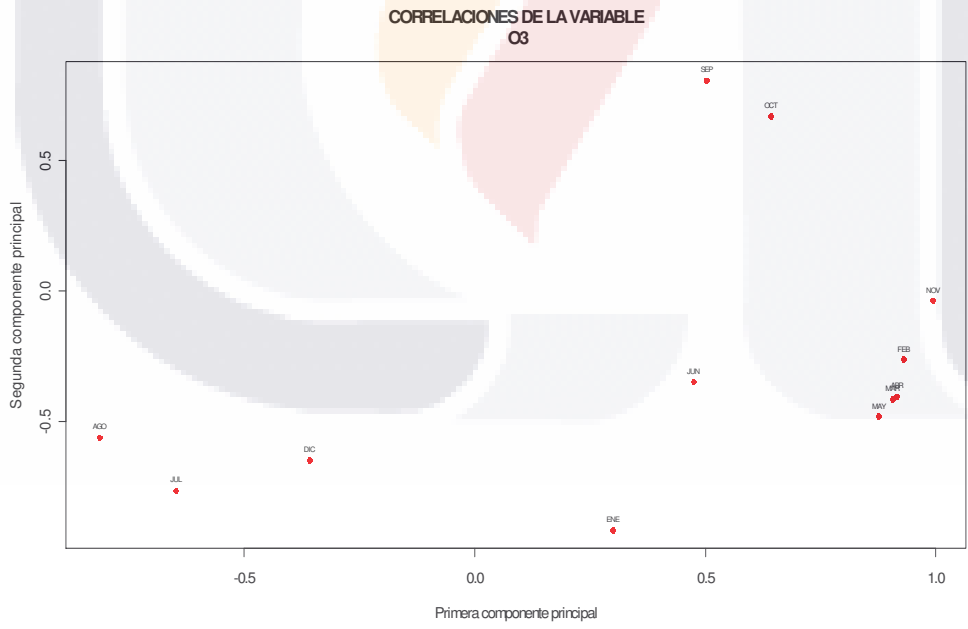


Figura 3.32: Correlaciones del Ozono a través de los meses

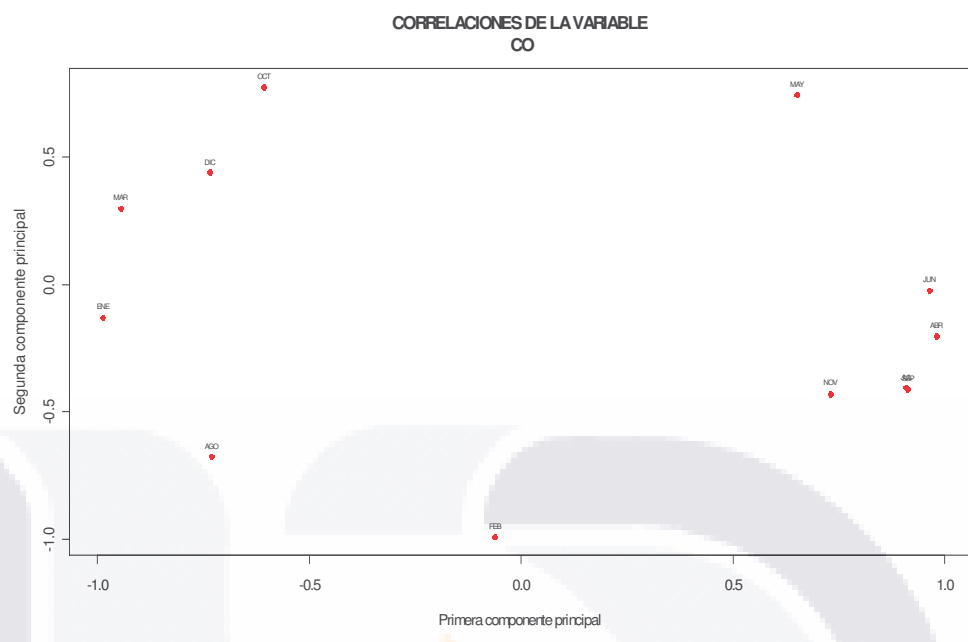


Figura 3.33: Correlaciones del Monóxido de Carbono a través de los meses

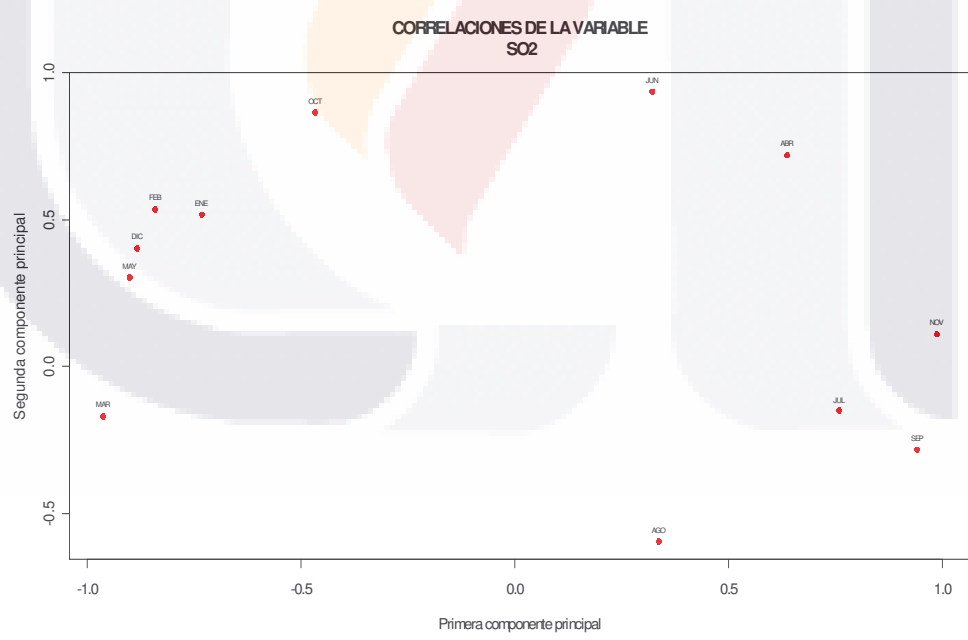


Figura 3.34: Correlaciones del Dióxido de Azufre a través de los meses

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	FEB	ENE	FEB	PORCENTAJE
<b>Celaya</b>	-0.148	0.279	0.833	1.518	-82.3
<b>Irapuato</b>	-0.527	-0.502	0.454	0.737	-62.2
<b>Salamanca</b>	-0.981	-1.239	0.000	0.000	0.0
<b>León</b>	1.655	1.462	2.636	2.700	-2.4

Figura 3.35: Diferencia en los puntajes de Enero y Febrero de 2006

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	MAR	ENE	MAR	PORCENTAJE
<b>Celaya</b>	-0.148	0.434	0.833	1.832	-120.0
<b>Irapuato</b>	-0.527	-0.353	0.454	1.045	-130.2
<b>Salamanca</b>	-0.981	-1.398	0.000	0.000	0.0
<b>León</b>	1.655	1.316	2.636	2.714	-3.0

Figura 3.36: Diferencia en los puntajes de Enero y Marzo de 2006

por lo que no se puede comparar un mes con otro. Una técnica utilizada para analizar los cambios por mes, es realizar una tabla tomando como base una ciudad que ha permanecido relativamente constante a los largo del tiempo. Así, se obtiene la diferencia entre el índice de contaminación de las ciudades y la mejor ciudad para 2 meses distintos, y posteriormente se obtiene el porcentaje en que cambió. La primeras 3 tablas generadas para las diferencias de los puntajes con respecto al puntaje mayor, se muestran en las figuras 3.35, 3.36 y 3.37.

Con estas tablas no es posible conocer respecto a cual variable se ha incrementado o reducido la distancia. Además, no es posible conocer el desarrollo de una misma ciudad a través del tiempo, ya que se compara en relación a otra.

Existen desventajas al realizar un análisis de componentes principales, ya que si se desean analizar simultáneamente varias matrices de datos, las variables pudieran estar correlacionadas de una matriz a otra, o no sería posible conocer la evolución de los individuos u observaciones. Para ello se utiliza una metodología llamada STATIS, la cual analiza datos numéricos con el objeto de conocer el comportamiento de las variables medidas en las ciudades de Guanajuato.

Para utilizar la metodología primero es necesario determinar el peso que tendrá cada individuo

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	ABR	ENE	ABR	PORCENTAJE
<b>Celaya</b>	-0.148	-0.843	0.833	0.000	100.0
<b>Irapuato</b>	-0.527	1.707	0.454	2.550	-461.6
<b>Salamanca</b>	-0.981	-0.404	0.000	0.439	0.0
<b>León</b>	1.655	-0.461	2.636	0.382	85.5

Figura 3.37: Diferencia en los puntajes de Enero y Abril de 2006



Enero	Celaya	Irapuato	Salamanca	León
Celaya	7.30	1.18	-4.44	-4.04
Irapuato	1.18	4.60	0.85	-6.63
Salamanca	-4.44	0.85	12.51	-8.92
León	-4.04	-6.63	-8.92	19.60

Figura 3.38: Matriz de Productos Escalares de Enero

Febrero	Celaya	Irapuato	Salamanca	León
Celaya	11.64	-1.12	-6.62	-3.89
Irapuato	-1.12	4.39	1.51	-4.78
Salamanca	-6.62	1.51	12.20	-7.09
León	-3.89	-4.78	-7.09	15.77

Figura 3.39: Matriz de Productos Escalares de Febrero

y cada arreglo. En nuestro caso, se toma un peso igual para cada mes, es decir, el peso de cada uno es de  $\frac{1}{12}$ , ya que no existe razón por la cual dar mayor importancia a un mes respecto al resto. De la misma manera, el peso que se utilizará para cada ciudad de Guanajuato será de  $\frac{1}{4}$ , ya que no existe diferencia entre una ciudad y otra.

### 3.3 Intraestructura e Interestructura

La matriz de productos escalares entre los individuos se obtiene por medio de la matriz de sumas de cuadrados y productos cruzados de  $Z$ . Las matrices de los primeros tres meses se muestran en las figuras 3.38, 3.39 y 3.40.

La Intraestructura de cada estudio  $W_k$ , se obtiene multiplicando la matriz de productos escalares por la matriz diagonal con los pesos de las observaciones, en nuestro caso,  $\frac{1}{4}$ . Las intraestructura de los primeros 3 meses se muestran en las figuras 3.41, 3.42 y 3.43.

Éstas matrices son relativamente estables a lo largo del tiempo, es decir, tienen normas  $\|W_k\|$  similares para cada uno de los arreglos a lo largo del tiempo (ver figura 3.44), y además, tienen productos escalares normados  $\left(\frac{W_k}{\|W_k\|}\right)$  cercanos a 1 (ver figura 3.45). La matriz de productos escalares normados es la matriz de los coeficientes de correlación vectorial ( $RV$ ).

De estos resultados obtenemos la Imagen Euclidiana de la figura 3.46, que representa la

Marzo	Celaya	Irapuato	Salamanca	León
Celaya	8.33	-3.14	-5.02	-0.17
Irapuato	-3.14	6.29	0.60	-3.75
Salamanca	-5.02	0.60	14.94	-10.52
León	-0.17	-3.75	-10.52	14.44

Figura 3.40: Matriz de Productos Escalares de Marzo

Enero	Celaya	Irapuato	Salamanca	León
Celaya	1.82	0.30	-1.11	-1.01
Irapuato	0.30	1.15	0.21	-1.66
Salamanca	-1.11	0.21	3.13	-2.23
León	-1.01	-1.66	-2.23	4.90

Figura 3.41: Intraestructura de Enero de 2006

Febrero	Celaya	Irapuato	Salamanca	León
Celaya	2.91	-0.28	-1.66	-0.97
Irapuato	-0.28	1.10	0.38	-1.20
Salamanca	-1.66	0.38	3.05	-1.77
León	-0.97	-1.20	-1.77	3.94

Figura 3.42: Intraestructura de Febrero de 2006

Marzo	Celaya	Irapuato	Salamanca	León
Celaya	2.08	-0.79	-1.26	-0.04
Irapuato	-0.79	1.57	0.15	-0.94
Salamanca	-1.26	0.15	3.73	-2.63
León	-0.04	-0.94	-2.63	3.61

Figura 3.43: Intraestructura de Marzo de 2006

	Normas $W_k$
1	7.66
2	7.18
3	7.34
4	7.16
5	7.25
6	6.93
7	6.92
8	6.68
9	7.11
10	6.62
11	6.93
12	7.25

Figura 3.44: Normas  $W_k$

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Enero	1	0.96	0.93	0.59	0.81	0.87	0.92	0.76	0.63	0.68	0.82	0.84
Febrero	0.96	1	0.95	0.64	0.89	0.88	0.92	0.77	0.72	0.80	0.89	0.84
Marzo	0.93	0.95	1	0.73	0.74	0.83	0.83	0.81	0.74	0.77	0.84	0.85
Abril	0.59	0.64	0.73	1	0.57	0.68	0.66	0.88	0.83	0.86	0.68	0.85
Mayo	0.81	0.89	0.74	0.57	1	0.90	0.94	0.77	0.77	0.86	0.94	0.68
Junio	0.87	0.88	0.83	0.68	0.90	1	0.98	0.93	0.76	0.86	0.91	0.71
Julio	0.92	0.92	0.83	0.66	0.94	0.98	1	0.88	0.75	0.83	0.91	0.77
Agosto	0.76	0.77	0.81	0.88	0.77	0.93	0.88	1	0.84	0.91	0.85	0.76
Septiembre	0.63	0.72	0.74	0.83	0.77	0.76	0.75	0.84	1	0.89	0.92	0.73
Octubre	0.68	0.80	0.77	0.86	0.86	0.86	0.83	0.91	0.89	1	0.88	0.75
Noviembre	0.82	0.89	0.84	0.68	0.94	0.91	0.91	0.85	0.92	0.88	1	0.73
Diciembre	0.84	0.84	0.85	0.85	0.68	0.71	0.77	0.76	0.73	0.75	0.73	1

Figura 3.45: Productos Escalares Normados *RV*

interestructura de los cuatro estudios analizados. La oposición entre enero y abril se presenta porque el producto escalar normado correspondiente es el más débil (0.59), y refleja los cambios importantes (relativamente) en los contaminantes ambientales y las variables meteorológicas. La cercanía entre los meses de mayo, junio y julio se explica por fuertes productos escalares normados que tienen (mayo-junio 0.90, mayo-julio 0.94, y junio-julio 0.98). Además, se observa que los grupos {Enero, Febrero, Marzo}, {Mayo, Junio, Julio} y {Octubre, Noviembre, Diciembre} son grupos estables, en el sentido que la estructura de correlaciones se mantiene en el tiempo.

La información contenida en la Imagen Euclidiana, junto con el hecho de obtener normas similares en cada uno de los arreglos, y productos normados cercanos a 1, son indicativos de una evolución cronológica de la contaminación de las ciudades. Además, es posible comparar la información de todos los meses, e interpretar las trayectorias encontradas.

### 3.4 Compromiso

Como los 12 estudios son comparables, se construye el compromiso de las ciudades del corredor industrial de Guanajuato, que se presenta en la figura 3.47. Su gráfica nos da información de la similitud o disimilitud de las ciudades en cuestión. En este sentido, si dos ciudades se encuentran cerca (como es el caso de Irapuato y Celaya), indica que tienen características similares, en términos de los contaminantes ambientales y las variables meteorológicas. Similarmente, si dos ciudades se encuentran alejadas (como es el caso de León y Salamanca), significa que las características medidas en ellas, difieren.

Como se puede ver en las figuras 3.48, 3.49 y 3.50, en términos generales con respecto al eje primario se tiene un contraste, por un lado con valores positivos el Monóxido de Carbono, el Dióxido de Azufre y la Presión Barométrica; con signo negativo el Ozono, el Dióxido de Nitrógeno, el Óxido de Nitrógeno, la Temperatura y la Radiación Solar, no obstante que en algunas variables el comportamiento cambia inclusive de signo, es decir, va de valores positivos grandes a valores negativos grandes, con respecto a las correlaciones con los compromisos.

En relación con el eje secundario, podemos observar un contraste de manera general, con valores positivos de las variables Humedad Relativa y Velocidad del viento, y valores negativos de Dióxido de Azufre.

Como las correlaciones con los compromisos cambian con respecto a los meses, la interpretación se complica, pues indica que la forma como interactúan los contaminantes ambientales y las variables meteorológicas, cambia a lo largo del año.

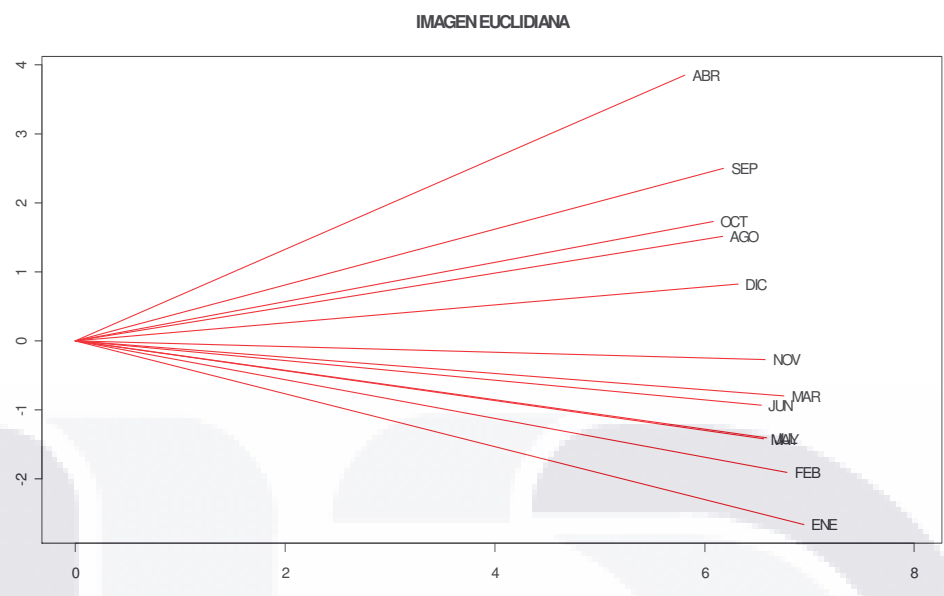


Figura 3.46: Imagen Euclidiana

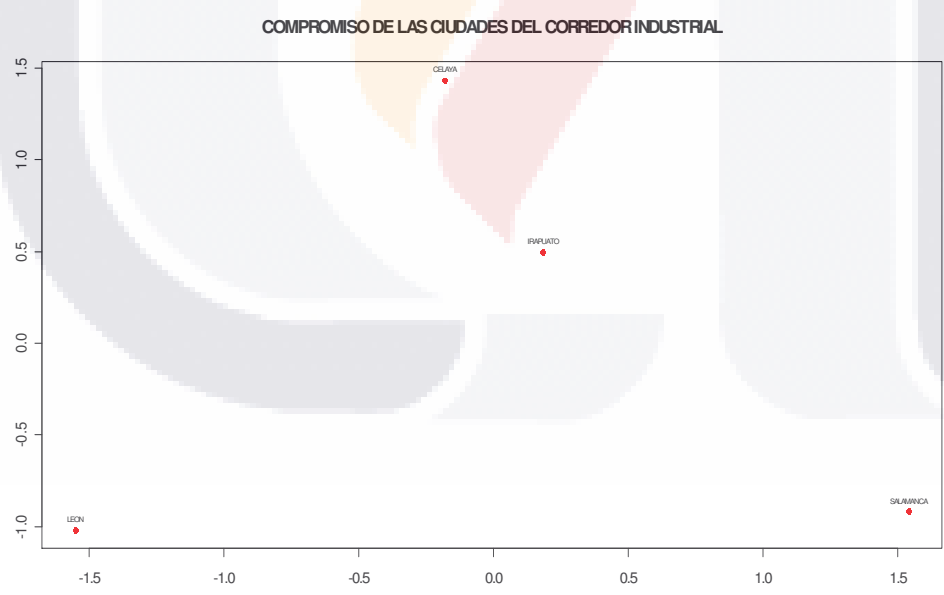


Figura 3.47: Compromiso de las Ciudades de Guanajuato

EJE PRINCIPAL												
VARIABLES	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
O3	-0.54	-0.86	-0.95	-0.28	-0.42	-0.18	-0.14	-0.16	-0.13	0.68	0.59	-0.39
CO	0.88	0.08	0.97	-0.07	0.76	-0.87	-0.90	-0.07	0.61	0.65	-0.05	0.51
SO2	0.80	0.84	0.94	0.68	0.22	0.24	-0.46	0.92	0.76	0.70	0.70	0.73
NO	-0.80	0.46	0.81	0.31	0.41	-0.69	-0.78	-0.89	-0.70	-0.04	-0.61	0.62
NO2	-0.91	-0.93	0.44	0.41	0.07	-0.73	-0.41	-0.62	-0.39	-0.12	0.01	0.66
NOX	-0.75	-0.91	0.80	0.41	0.45	-0.67	-0.72	-0.90	-0.41	0.27	-0.58	0.62
WS	-0.08	-0.08	0.66	-0.11	-0.08	-0.09	-0.10	-0.27	-0.14	-0.31	0.89	0.11
TEMP	-0.90	-0.72	-0.81	-0.85	-0.96	-0.84	0.68	0.69	0.89	0.79	0.73	-0.78
HR	0.43	0.28	-0.01	0.68	0.20	0.04	-0.05	-0.03	0.34	0.39	0.54	-0.12
PBAR	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.70	0.70	0.71	0.88	0.88
RADSOL	-0.99	-0.92	-0.72	-0.56	-0.69	-0.89	-0.91	-0.19	-0.13	-0.15	-0.49	-0.94

Figura 3.48: Correlación del eje principal con el Compromiso

EJE SECUNDARIO												
VARIABLES	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
O3	0.79	0.15	-0.31	0.19	0.91	-0.04	0.98	0.16	-0.28	0.24	-0.81	0.28
CO	0.42	0.94	0.23	0.17	0.57	-0.48	-0.42	0.29	-0.80	-0.75	-0.87	-0.44
SO2	-0.29	-0.49	-0.35	-0.41	-0.96	-0.97	-0.55	-0.16	-0.64	-0.71	-0.71	-0.23
NO	-0.51	0.79	-0.22	0.17	0.90	-0.52	-0.57	-0.15	0.18	-0.23	-0.22	0.19
NO2	-0.22	-0.24	0.10	0.09	0.72	-0.67	-0.90	-0.72	0.60	-0.86	0.87	0.37
NOX	-0.66	-0.40	-0.29	0.03	0.79	-0.71	-0.68	-0.22	0.85	-0.31	0.79	0.19
WS	0.95	0.90	0.31	0.90	0.90	0.90	0.90	0.79	0.78	0.71	-0.42	0.93
TEMP	-0.40	-0.30	-0.58	-0.17	-0.29	-0.38	0.21	0.57	-0.26	0.59	0.67	-0.62
HR	0.86	0.96	0.93	0.54	0.80	0.60	0.70	0.92	0.89	0.87	0.79	0.93
PBAR	-0.48	-0.48	-0.47	-0.47	-0.47	-0.47	-0.48	-0.49	-0.49	-0.49	-0.45	-0.45
RADSOL	0.07	0.32	0.64	0.38	0.64	0.46	0.40	-0.17	-0.27	-0.26	-0.33	-0.20

Figura 3.49: Correlaciones del eje secundario con el Compromiso

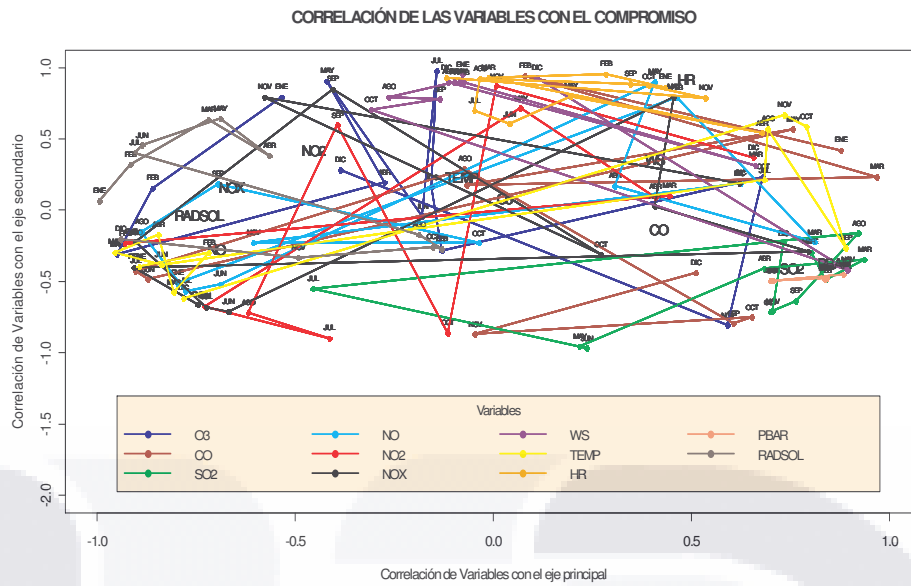


Figura 3.50: Correlación de las Variables con el Compromiso

A diferencia de otro tipo de aplicaciones donde es posible nombrar a los ejes principales a través de las variables que se asocian a ellos, en este caso es difícil darle nombre a los dos ejes, por una parte porque incluyen tanto variables meteorológicas como contaminantes atmosféricos, y por otro lado, porque el comportamiento de las relaciones cambia a lo largo del año.

Realizando un Análisis de Componentes Principales a Sdelta, se obtiene una tabla resumen de los eigenvalores (ver figura 3.51). La gráfica de la ladera correspondiente se muestra en la figura 3.52.

Como se puede apreciar, con dos valores propios se logra acumular el 90% de la variabilidad medida por Sdelta. En otras palabras, con las dos primeras dimensiones podemos representar de manera adecuada la información. La gráfica de ladera sugiere el uso de un solo componente principal, pero como ya mencionamos, usaremos los dos primeros.

### 3.5 Trayectorias

Como se ha mencionado, se puede caracterizar de manera general a cada una de las ciudades, dependiendo de su posición en la gráfica de compromisos de los individuos. Pero como se indicó, nos interesa conocer los cambios de las ciudades en el tiempo, por ésta razón se construyen las trayectorias por ciudad.

Debemos recordar que las trayectorias se interpretan respecto a una trayectoria promedio que se encontraría en el origen de la gráfica. En la figura 3.53 se muestran los puntos correspondientes al mes de enero para las cuatro ciudades. Ésta gráfica se puede analizar de manera similar a la gráfica de compromisos de las ciudades.

La figura 3.54 muestra las trayectorias de las ciudades en los meses enero y febrero. Como se puede apreciar en ésta gráfica, un alejamiento con respecto al origen representa un alejamiento de una ciudad promedio, en término de las variables en cuestión. En consecuencia, las ciudades

Componente	Eigenvalor	Porcentaje	% Acumulado
1	499.57	82.8	82.8
2	43.43	7.2	90
3	31.97	5.3	95.3
4	13.60	2.3	97.6
5	9.03	1.5	99.1
6	5.69	0.9	100
7	0	0	100
8	0	0	100
9	0	0	100
10	0	0	100
11	0	0	100
12	0	0	100

Figura 3.51: Eigenvalores de Sdelta

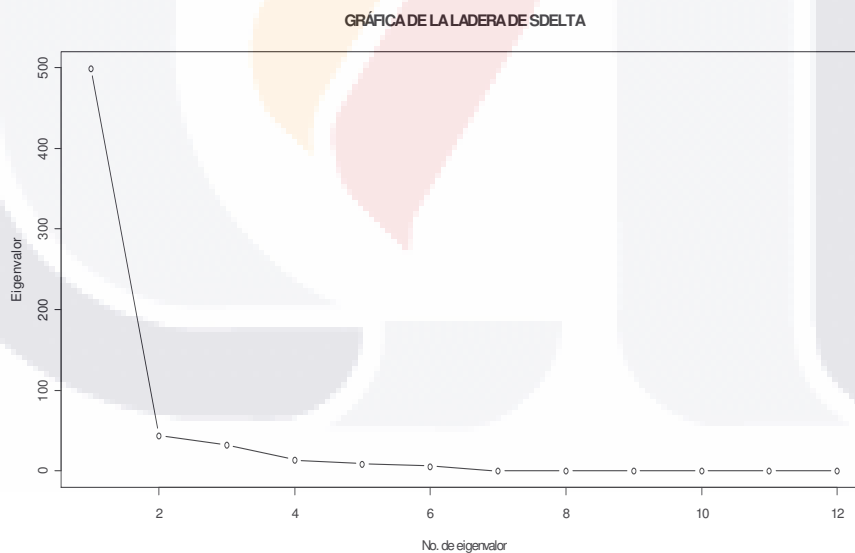


Figura 3.52: Gráfica de Ladera de Sdelta

TRAYECTORIAS DE LAS CIUDADES DEL CORREDOR INDUSTRIAL  
DE ENE A ENE DE 2006

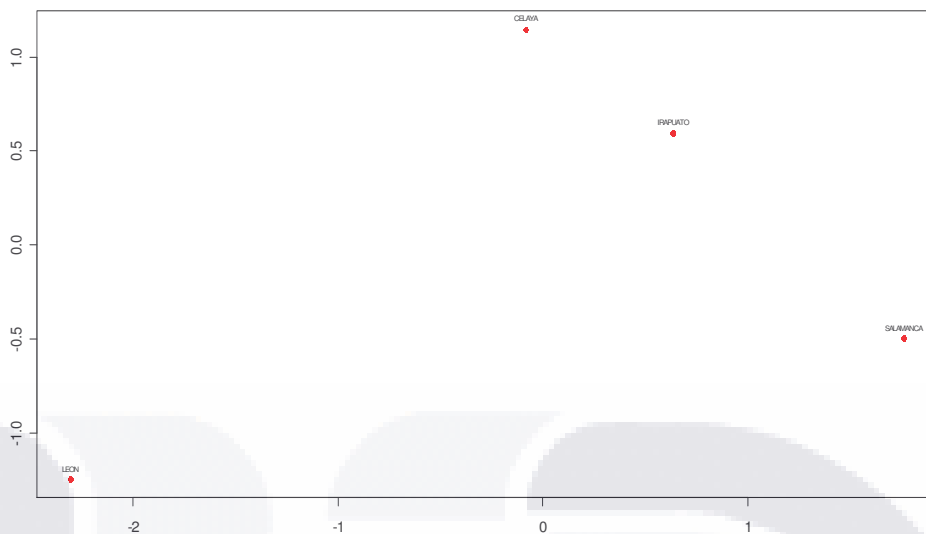


Figura 3.53: Trayectorias de las ciudades en el mes de Enero de 2006

de Irapuato, Salamanca y León, en febrero se comportaron más como una ciudad promedio, a diferencia de Celaya que se aleja de este comportamiento en febrero.

En la figura 3.55 se muestra la trayectoria a lo largo de los doce meses, y como se aprecia de manera general, el comportamiento de las ciudades es similar a lo largo del año, es decir, no hay cambios severos, de hecho, se puede apreciar que están en zonas bien definidas. Las ciudades que tienden a tener un comportamiento similar son Celaya e Irapuato. Salamanca y León se alejan de ese patrón. Todas las ciudades en abril tienen un comportamiento próximo al comportamiento promedio.

De manera muy general, se podría decir que Salamanca a lo largo del año presenta valores altos de Dióxido de Azufre y de Monóxido de Carbono, y se registraron en 2006 valores altos de la Presión Barométrica, además, de valores bajos de Ozono, Dióxido de Nitrógeno, Óxido de Nitrógeno, así como registros bajos de Temperatura y Radiación Solar.

La ciudad de León por el contrario, presenta niveles bajos de Dióxido de Azufre, Monóxido de Carbono y Presión Barométrica, pero niveles altos de Dióxido de Nitrógeno, Óxido de Nitrógeno y Ozono; además de niveles altos de Temperatura y Radiación Solar.

La ciudad de Celaya se caracterizaría por niveles bajos de Dióxido de Azufre y observaciones altas de Humedad Relativa, Velocidad del viento y Ozono en algunos meses. En relación con las otras variables no tendría un comportamiento claro.

La ciudad de Irapuato en términos generales, se comportaría más como la ciudad promedio, es decir, no tiene un comportamiento específico con algún contaminante ambiental o variable meteorológica.



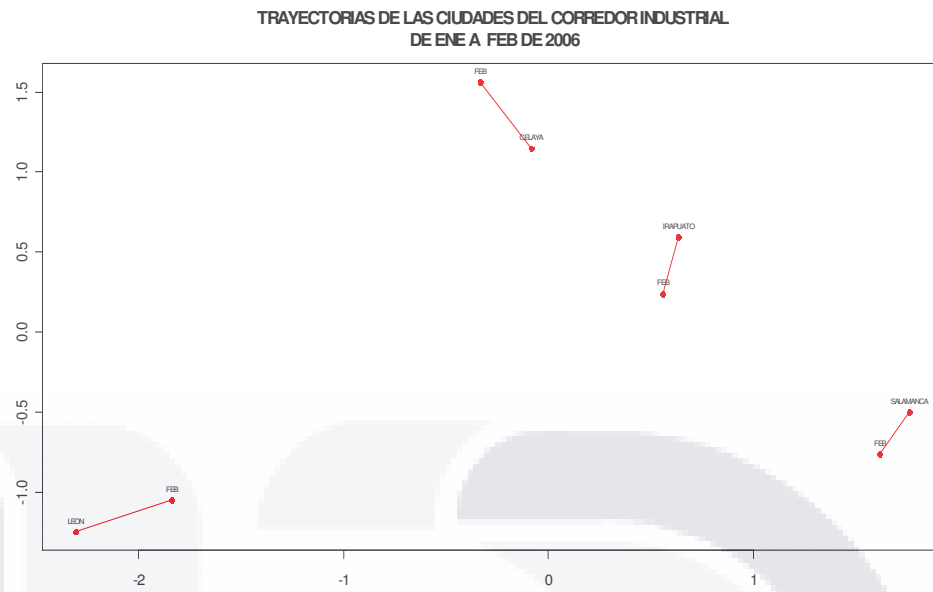


Figura 3.54: Trayectorias de las ciudades de Enero a Febrero de 2006

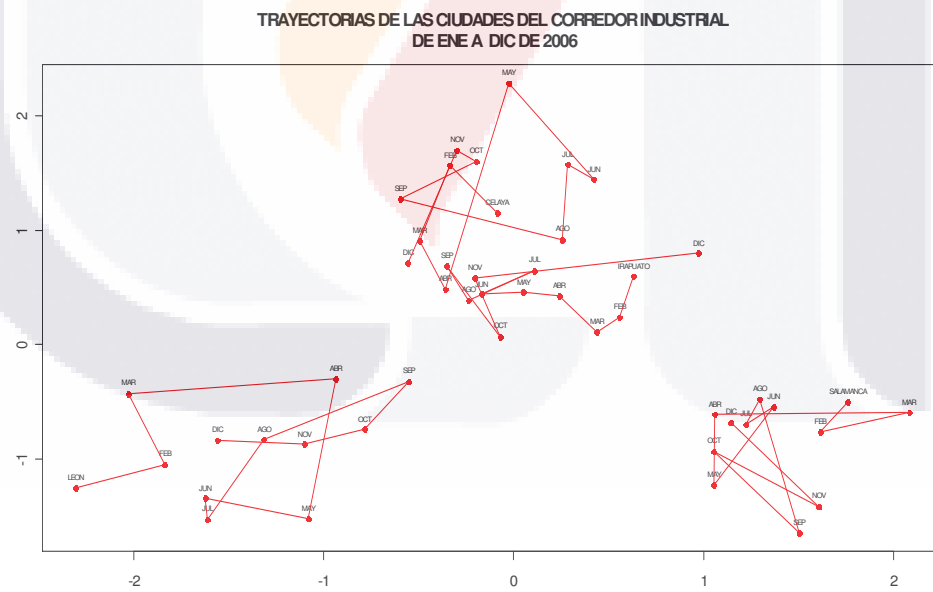


Figura 3.55: Trayectorias de las Ciudades de Guanajuato

TESIS TESIS TESIS TESIS TESIS



TESIS TESIS TESIS TESIS TESIS

## Conclusiones

El estudio de la contaminación ambiental es de vital importancia y se ha venido estudiando usando diferentes metodologías, tanto de manera multivariada, mediante el uso de Análisis de Componentes Principales por ejemplo, como de forma longitudinal univariada. El fenómeno de la contaminación es muy complejo, por una parte por la dinámica de la contaminación, y por otra, por la falta de información confiable.

Para el análisis descriptivo multivariado existen varias alternativas, la más común es el uso del Análisis de Componentes Principales comunes, pero supone que los individuos son independientes, de forma que sería imposible introducir el hecho que las observaciones son de los mismos sujetos y las mismas variables, en diferentes tiempos. Por ésta razón se utiliza la Metodología STATIS. Como se ha mencionado, ésta metodología permite estudiar información de diferentes esquemas, dependiendo de como se recolectaron los datos, es decir, si se recolectan medidas de los mismos individuos en diferentes ocasiones, entonces se tendrá un arreglo para cada ocasión; o si las mismas variables se miden en diferentes grupos de individuos, entonces se tendrá un arreglo que corresponda a cada grupo distinto de individuos. En nuestro estudio utilizamos el primer esquema.

Al llevar a cabo el análisis, observamos que la información de las variables medidas cambia a lo largo del tiempo, es decir, el comportamiento no es el mismo.

El tipo de contaminación de Salamanca es en buen medida originada por la actividad petroquímica de la zona, lo que origina niveles altos principalmente de Dióxido de Azufre, a diferencia de León, donde se caracteriza por la industria de la curtiduría, que se caracteriza por niveles altos de Óxidos de Nitrógeno y Dióxido de Nitrógeno.

Como se mencionó, para varias variables el comportamiento a lo largo del año cambia, de forma que algunas ciudades pueden tener valores altos de alguna variable para un mes específico y para el siguiente tener valores pequeños.

En general, la Metodología STATIS nos proporcionó información importante del estudio de la contaminación ambiental a través de los meses. Para futuros trabajos se podría utilizar esta metodología con otros estadísticos, como por ejemplo el percentil 95 o la mediana.

TESIS TESIS TESIS TESIS TESIS



TESIS TESIS TESIS TESIS TESIS

## Apéndice A

# Análisis de Datos y Gráficos en R

R es un sistema para análisis estadísticos y gráficos creado por Ross Ihaka y Robert Gentleman del Departamento de Estadística de la Universidad de Auckland (ver figura A.1). Tiene una doble naturaleza, como programa y como lenguaje de programación, y es considerado como un dialecto del lenguaje **S** creado por los Laboratorios AT&T Bell. R se distribuye gratuitamente bajo los términos de la **GNU** (*General Public Licence*), y está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux. Su desarrollo y distribución son llevados a cabo por varios estadísticos conocidos como el Grupo Nuclear de Desarrollo de R [PEAJ2003].

R está disponible en varias formas, el código fuente escrito principalmente en C (y algunas rutinas en Fortran), esencialmente para máquinas Unix y Linux, o como archivos binarios precompilados para Windows, Linux (Debian, Mandrake, RedHat, SuSe), Macintosh y Alpha Unix.

Los archivos necesarios para instalar R, ya sea desde las fuentes o binarios precompilados, se distribuyen desde el sitio de internet **CRAN** (*Comprehensive R Archive Network*) junto con las instrucciones de instalación, [RPSC2009].

### A.1 Características

R posee muchas características que hacen que sobresalga de varios softwares estadísticos. Algunas de éstas características se mencionan a continuación.

R es un lenguaje *Orientado a Objetos*, bajo este término se esconde su simplicidad y flexibilidad. Es un lenguaje interpretado (como Java) y no compilado (como C, C++, Fortran, Pascal, etc.), lo cual significa que los comandos escritos en el teclado son ejecutados directamente sin necesidad de construir ejecutables. En un *Lenguaje Orientado a Objetos* las variables, datos, funciones, resultados, etc., se guardan en la memoria activa del computador en forma de *objetos* con un nombre específico, y el usuario puede modificar o manipular estos objetos con operadores (aritméticos, lógicos, y comparativos), y funciones (que a su vez son objetos).



Figura A.1: Logotipo de R

R cuenta con muchas funciones para análisis estadísticos y gráficos; estos últimos pueden ser visualizados de manera inmediata en su propia ventana y ser guardados en varios formatos (*jpg, png, bmp, ps, pdf, emf, pictex, xfig*; los formatos disponibles dependen del sistema operativo). Los resultados de análisis estadísticos se muestran en la pantalla, y algunos resultados intermedios (como valores P, coeficientes de regresión, residuales, etc.) se pueden guardar, exportar a un archivo, o ser utilizados en análisis posteriores.

Este lenguaje permite al usuario, programar ciclos para analizar conjuntos sucesivos de datos. También es posible combinar en un solo programa diferentes funciones estadísticas para realizar análisis más complejos. Hay disponible un gran número de programas escritos para **S** y que pueden ser utilizados directamente con R.

Una de las características más sobresalientes de R es su enorme flexibilidad. Mientras que programas clásicos muestran directamente los resultados de un análisis, R guarda estos resultados como un objeto, de tal manera que se puede hacer un análisis sin necesidad de mostrar su resultado inmediatamente, el usuario puede extraer solo aquella parte de los resultados que le interesa.

R ofrece una increíble variedad de gráficos. Cada función gráfica en R tiene un número de opciones permitiendo una gran producción de gráficos. Particularmente, el resultado de una función gráfica no puede ser asignado a un objeto, sino que es enviado a un dispositivo gráfico. Un dispositivo gráfico es una ventana gráfica o un archivo.

Existen dos tipos de funciones gráficas: las funciones de graficación de *alto nivel* que crean una nueva gráfica, y las funciones de graficación de *bajo nivel* que agregan elementos a una gráfica ya existente.

Dos de las principales características por la que se utilizó R en este estudio fueron, primero, la simplicidad en el manejo de arreglo de datos, ya que por medio de un *data frame* se agruparon variables numéricas y categóricas, de las cuales se obtuvo un arreglo de datos tridimensional, es decir, en los renglones las observaciones, en las columnas las variables y en las capas los tiempos. Segundo, las gráficas deseadas en éste estudio, se programan de tal manera que R las realiza.

## A.2 Literatura Adicional

R trae 5 manuales en inglés que se instalan por defecto en *R HOME/doc/manual/* (donde R HOME se ubica donde R está instalado). Los manuales son:

1. An Introduction to R: [R-intro.pdf].
2. R Installation and Administration: [R-admin.pdf].
3. R Data Import/Export: [R-data.pdf].
4. Writing R Extensions: [R-exts.pdf].
5. R Language Definition: [R-lang.pdf].

Los archivos pueden estar en diferentes formatos (*pdf, html, texi, . . .*) dependiendo del tipo de instalación.

El sitio **CRAN** y la página web de R contienen varios documentos, recursos bibliográficos y enlaces a otros sitios. También se puede encontrar allí una lista de publicaciones (libros y artículos) sobre R y métodos estadísticos en general, y algunos documentos y tutoriales escritos por usuarios de R.

## Apéndice B

# Programas en R

Se presentan a continuación los programas en R que obtienen los resultados y gráficas de la aplicación de la Metodología STATIS a los datos de contaminación ambiental y a las variables meteorológicas.

### B.1 Lectura y Arreglo de Datos

En esta sección se encuentra el programa que lee el archivo de los datos de contaminación guardado con el software SPSS. Después de convertirlos en un *data frame* solo se utilizan 6 contaminantes ambientales, 6 variables meteorológicas y 9 estaciones de monitoreo (la explicación del por qué se encuentra en el capítulo ??). Con dicha información se realiza un arreglo de datos de 3 entradas: las observaciones son las 4 ciudades del corredor industrial, las columnas son las 12 variables medidas y las capas son los 12 meses del año.

*Nota.* El archivo *todas.sav* deberá estar en el directorio donde se ejecutará el programa.

#### LECTURA DE LA INFORMACIÓN Y CONVERSIÓN A DATA FRAME

```
rm(list=ls())
library(foreign)
lista=read.spss("todas.sav")

MES=as.factor(lista$MES)
DIA=as.factor(lista$DIA)
HORA=as.factor(lista$HORA)
COLUMNA=lista$COLUMNA
O3=lista$O3
PM10=lista$PM10
CO=lista$CO
SO2=lista$SO2
NO=lista$NO
NO2=lista$NO2
NOX=lista$NOX
CH4=lista$CH4
NMHC=lista$NMHC
THC=lista$THC
WS=lista$WS
```

```

WD=lista$WD
TEMP=lista$TEMP
HR=lista$HR
PBAR=lista$PBAR
RADSOL=lista$RADSOL
TRIMESTRE=as.factor(lista$Trimestre)
O3_2=lista$O3_2
CO_2=lista$CO_2
SO2_2=lista$SO2_2
ESTACION=lista$estacion
CIUDAD=lista$CIUDAD

```

```

contaminacion=data.frame(MES, DIA, HORA, COLUMNA, O3, PM10, CO, SO2,
                          NO, NO2, NOX, CH4, NMHC, THC, WS, WD, TEMP, HR, PBAR,
                          RADSOL, TRIMESTRE, O3_2, CO_2, SO2_2, ESTACION, CIUDAD)
str(contaminacion)

```

#### **ADICIÓN DE LA VARIABLE COMPUESTA TIEMPO (MES+ DIA+HORA)**

```

attach(contaminacion)
TIEMPO=cbind(MES, DIA, HORA)
TIEMPO=paste(MES, DIA, HORA)
TIEMPO=as.factor(paste(sprintf("%02.0f", MES), sprintf("%02.0f", DIA),
                          sprintf("%02.0f", HORA), sep=""))
detach()
contaminacion=cbind(contaminacion, TIEMPO)

```

#### **CAMBIO DE VALORES NEGATIVOS A NA**

```

contaminacion[c(5:20, 22:24)][contaminacion[c(5:20, 22:24)] < 0]=NA

```

#### **ELIMINACIÓN DE LA ESTACIÓN SAL\_DIF**

```

paso=contaminacion
paso=paso[paso$ESTACION!="SAL DIF",]
temp=as.matrix(paso$ESTACION)
paso$ESTACION=as.factor(temp)
contaminacion=paso

```

#### **ARREGLO DE DATOS CON EL PORCENTAJE DE DATOS CORRECTOS**

```

mina<-function(zz)
{
  round(sum(!is.na(zz))/length(zz)*100, 1)
}

matrizpaso3=aggregate(contaminacion[, c(5:20, 22:24)],
                      list(lugar=contaminacion$ESTACION, tpo=contaminacion$MES), mina)
matriz1=as.matrix(matrizpaso3[, -(1:2)])
matriz2=array(t(matriz1), dim=c(ncol(matriz1), length(levels(matrizpaso3[, 1])),
                             length(levels(matrizpaso3[, 2]))))
dimnames(matriz2)=list(colnames(matrizpaso3[, -(1:2)]),

```



```

        levels(matrizpaso3[, 1]), levels(matrizpaso3[, 2]))
correctos=aperm(matriz2, c(3, 1, 2))
correctos
    
```

### ELIMINACIÓN DE LAS VARIABLES PM10, CH4, NMHC, THC Y WD

```

paso=contaminacion[-c(6, 12:14,165)]
contaminacion=paso
    
```

### ARREGLO DE DATOS CON EL MÁXIMO

```

maxa<-function(zz)
{
  x=max(zz, na.rm=TRUE)
  if (!is.finite(x))
    x=NA
  return(x)
}

matrizpaso1=aggregate(contaminacion[, 5:15], list(lugar=contaminacion$CIUDAD,
  tpo=contaminacion$TIEMPO), maxa)
matriz1=as.matrix(matrizpaso1[, -(1:2)])
matriz2=array(t(matriz1), dim=c(ncol(matriz1),
  length(levels(matrizpaso1[, 1])), length(levels(matrizpaso1[, 2]))))
dimnames(matriz2)=list(colnames(matrizpaso1[, -(1:2)]),
  levels(matrizpaso1[, 1]), levels(matrizpaso1[, 2]))
datos=aperm(matriz2, c(3, 1, 2))
    
```

### GRÁFICAS PARA LOS CONTAMINANTES

```

pdf(file="graficas.pdf", paper="letter")
m=matrix(c(1, 2), 2, 1)
layout(m)
for (i in 1:11)
{
  boxplot(as.data.frame(datos[, i, ]), main=paste("MEDICIONES DE ",
    dimnames(contaminacion[i+4])[2], " POR CIUDAD"),
    xlab="CIUDAD", ylab=dimnames(contaminacion[i+4])[2])
  for (j in 1:4)
  {
    if (is.na(summary(datos[, i, j])[7]) || (summary(datos[, i, j])[7]!=8760))
    {
      hist(datos[, i, j], main=paste(dimnames(contaminacion[i+4])[2],
        " EN ", levels(contaminacion$CIUDAD)[j]),
        xlab=dimnames(contaminacion[i+4])[2],
        ylab="Num. de Mediciones")
      plot(datos[, i, j], main=paste(dimnames(contaminacion[i+4])[2],
        " EN ", levels(contaminacion$CIUDAD)[j]),
        xlab="Orden de los datos",
        ylab=dimnames(contaminacion[i+4])[2])
    }
  }
}
    
```

```

    }
  }
}
dev.off()

```

### IMPUTACIÓN DE LAS VARIABLES NO, NO2, NOX

```

posicion=which(is.na(contaminacion$NO)==T & is.na(contaminacion$NO2)==F &
  is.na(contaminacion$NOX)==F)
contaminacion$NO[posicion]=abs(contaminacion$NOX[posicion]-
  contaminacion$NO2[posicion])
posicion=which(is.na(contaminacion$NO)==F & is.na(contaminacion$NO2)==T &
  is.na(contaminacion$NOX)==F)
contaminacion$NO2[posicion]=abs(contaminacion$NOX[posicion]-
  contaminacion$NO[posicion])
posicion=which(is.na(contaminacion$NO)==F & is.na(contaminacion$NO2)==F &
  is.na(contaminacion$NOX)==T)
contaminacion$NOX[posicion]=contaminacion$NO[posicion]+
  contaminacion$NO2[posicion]

```

### IMPUTACIÓN DEL MÁXIMO

```

todo=data.frame()
for (k in 1:4)
{
  corte=data.frame(CIUDAD=rep(dimnames(datos)[[3]][k], 8760),
    datos[, , k], TRIMESTRE=as.factor((as.numeric(substr(
      dimnames(datos)[[1]], 1, 2)) + 2) %/% 3),
    MES=as.factor((as.numeric(substr(dimnames(datos)[[1]], 1, 2)))),
    TIEMPO=dimnames(datos)[[1]])
  todo=rbind(todo, corte)
}

matrizpaso2=aggregate(todo[, 2:12],
list(lugar=todo$CIUDAD, tpo=todo$MES), maxa)
matriz1=as.matrix(matrizpaso2[, -(1:2)])
matriz2=array(t(matriz1), dim=c(ncol(matriz1),
  length(levels(matrizpaso2[, 1])), length(levels(matrizpaso2[, 2]))))
dimnames(matriz2)=list(colnames(matrizpaso2[, -(1:2)]),
  levels(matrizpaso2[, 1]), levels(matrizpaso2[, 2]))
datos2=aperm(matriz2,c(3, 1, 2))
dimnames(datos2)[[1]]=c("ENE", "FEB", "MAR", "ABR", "MAY",
  "JUN", "JUL", "AGO", "SEP", "OCT", "NOV", "DIC")

for (l in 1:4)
  for (m in 1:11)
    for (p in 1:12)
      if (datos2[p, m, l]==0 || is.na(datos2[p, m, l]))
        datos2[p, m, l]=datos2[p-1, m, l]

```

**ARREGLO DE DATOS PARA APLICAR METODOLOGÍA STATIS**

```
datos=aperm(datos2, c(3, 2, 1))
save(datos, file ="arreglo_cont.Rdata")
```

**B.2 Índices de Contaminación**

En esta sección se obtienen las matrices de estandarización ( $Z$ ), y de Varianzas - Covarianzas ( $V$ ) de los datos. Se aplica un análisis de componentes principales a la matriz  $V$  y se obtienen los índices de contaminación. Además se gráficán tanto los eigenvalores de  $V$ , las correlaciones entre variables y los índices de contaminación.

**LECTURA DEL ARREGLO DE DATOS**

```
rm(list=ls())
load("arreglo_cont.Rdata")
```

**DECLARACIÓN DE VARIABLES**

```
num.est = 12
num.vars = 11
num.obs = 4
peso.obs = 1/num.obs
peso.est = 1/num.est
tiempo = dimnames(datos)[[3]]
ciudad = dimnames(datos)[[1]]
vars = dimnames(datos)[[2]]
```

**MATRICES DE ESTANDARIZACIÓN Y DE VARIANZAS**

```
Z=array(0, dim=c(num.obs, num.vars, num.est))
R=array(0, dim=c(num.vars, num.vars, num.est))
for(i in 1:num.est)
  Z[, , i]=scale(datos[, , i],scale=sd(datos[, , i])*sqrt((num.obs-1)/num.obs))
for(i in 1:num.est)
  R[, , i]=cor(Z[, , i])
```

**ANÁLISIS DE COMPONENTES PRINCIPALES DE R**

```
pca=array(0, dim=c(num.obs, num.vars, num.est))
eigen.vec=array(0, dim=c(num.vars, num.vars, num.est))
eigen.val=matrix(0, num.vars, num.est)
for(i in 1:num.est)
{
  comp=eigen(Restdd[, , i], EISPACK=TRUE)
  eigen.vec[, , i]=as.matrix(comp$vectors)
  eigen.val[, i]=as.matrix(comp$values)
  pca[, , i]=Z[, , i]%*%eigen.vec[, , i]
}
```

**TABLA RESUMEN DE EIGENVALORES POR MES**

```

tab=array(0, dim=c(num.vars, 4, num.vars))
for(i in 1:num.est)
{
  colnames(tab)=c("Componente", "Eigenvalor", "Porcentaje", "% Acumulado")
  tab[, 1, i]=1:num.vars
  tab[, 2, i]=round(eigen.val[, i], digits=3)
  tab[, 3, i]=round((eigen.val[, i]/sum(eigen.val[, i]))*100, digits=1)
  tab[, 4, i]=round(cumsum(tab[, 3, i]), digits=1)
}

```

#### GRÁFICAS DE LA LADERA POR MES

```

for(i in 1:num.est)
{
  x11()
  plot(tab[, 1, i], tab[, 2, i], main=paste("GRÁFICA DE LA
  LADERA DEL MES DE\n", tiempo[i], "DE 2006"),
  xlab="No. de eigenvalor", ylab="Eigenvalor", type="b")
}

```

#### GRÁFICAS DE CADA CIUDAD

```

for(i in 1:num.obs)
{
  x11()
  plot(tempo1[1:12,i], tempo1[13:24,i], pch=16, type="p", col="red",
  main=paste("PUNTAJES DE LOS PRIMEROS 2 COMPONENTES EN\n", ciu-
dad[i]),
  xlab="Primera componente principal", ylab="Segunda componente principal")
  text(tempo1[1:12,i], tempo1[13:24,i], cex=.5, pos=3, labels=tiempo)
}

```

#### GRÁFICAS DE CORRELACIONES

```

tempo2=array(0, dim=c(2*num.est, num.vars))
for(i in 1:num.est)
{
  x11()
  x=eigen.vec[,1,i]*sqrt(eigen.val[1,i])
  y=eigen.vec[,2,i]*sqrt(eigen.val[2,i])
  plot(x, y, pch=16, type="p", col="red",
  main=paste("CORRELACIÓN DE VARIABLES EN EL\n", "MES DE ",
  tiempo[i], "DE 2006"),
  xlab="Correlación de Variables con el eje principal",
  ylab="Correlación de Variables con el eje secundario")
  text(x, y, cex=.5, pos=3, labels=vars)
  tempo2[i,]=x
  tempo2[i+12,]=y
}

```

#### GRÁFICAS DE CORRELACIONES PARA CADA VARIABLE

```
for(i in 1:num.vars)
{
  x11()
  plot(tempo2[1:12,i], tempo2[13:24,i], pch=16, type="p", col="red",
  main=paste("CORRELACIONES DE LA VARIABLE\n", vars[i]),
  xlab="Primera componente principal", ylab="Segunda componente principal")
  text(tempo2[1:12,i], tempo2[13:24,i], cex=.5, pos=3, labels=tiempo)
}
```

**PUNTAJES DE LOS 2 PRIMEROS COMPONENTES PRINCIPALES**

```
IC1=matrix(0, num.obs, num.est)
IC2=matrix(0, num.obs, num.est)
for(i in 1:num.est)
{
  IC1[,i]=pca[,1,i]/sqrt(eigen.val[1,i])
  IC2[,i]=pca[,2,i]/sqrt(eigen.val[2,i])
}
```

**TABLA DE DIFERENCIAS EN LOS ÍNDICES DE CONTAMINACIÓN RESPECTO AL ÍNDICE MAYOR**

```
reduccion =c("IC_1", "IC_2", "Dist_1", "Dist_2", "%_Reducción")
tabla =array(0, dim=c(num.obs, 5, choose(num.est, 2)))
nCr =combn(num.est, 2)
pareja =c(1:choose(num.est, 2))
dimnames(tabla)=list(ciudad, reduccion, pareja)
for(i in 1:choose(num.est, 2))
{
  tabla[, 1, i]=IC1[,nCr[1, i]]
  tabla[, 2, i]=IC1[,nCr[2, i]]
  tabla[, 3, i]=tabla[, 1, i]-min(tabla[, 1, i])
  tabla[, 4, i]=tabla[, 2, i]-min(tabla[, 2, i])
  tabla[, 5, i]=(tabla[, 3, i]-tabla[, 4, i])/tabla[, 3, i]*100
}
```

**B.3 Metodología STATIS**

En ésta sección se calculan la Intraestructura e Interestructura de los datos, se obtienen los productos escalares entre objetos  $Wk$ , 2 análisis de componentes principales, uno para  $SDelta$  y otro para  $WD$ . Además se realizan las gráficas de la Imágen Euclidiana, los compromisos de las ciudades y las gráficas de las trayectorias.

**PRODUCTOS ESCALARES ENTRE LOS INDIVIDUOS**

```
mpe=array(0, dim=c(num.obs, num.obs, num.est))
for(i in 1:num.est)
mpe[, , i]=Z[, , i]*%*%t(Z[, , i])
```

### INTRAESTRUCTURA DE CADA ESTUDIO

```
D=diag(peso.obs, num.obs, num.obs)
Wk=array(0, dim=c(num.obs, num.obs, num.est))
for(i in 1:num.est)
Wk[, , i]=mpe[, , i]%*%D
```

### INTERESTRUCTURA

```
S=matrix(0, num.est, num.est)
for(i in 1:num.est)
for(j in 1:num.est)
S[i, j]=sum(diag(Wk[, , i]%*%Wk[, , j]))
```

### COEFICIENTES DE CORRELACIÓN VECTORIAL RV

```
norma.Wk=sqrt(diag(S))
RV=diag(1/norma.Wk)%*%S%*%diag(1/norma.Wk)
```

### ANÁLISIS DE COMPONENTES PRINCIPALES DE SDELTA

```
delta=diag(1, num.est, num.est)
Sdelta=S%*%delta
comp.Sdelta=eigen(Sdelta, EISPACK=TRUE)
vec.Sdelta=as.matrix(comp.Sdelta$vector)
val.Sdelta=comp.Sdelta$values
pca.Sdelta=vec.Sdelta[, 1:2]%*%diag(sqrt(as.vector(val.Sdelta[1:2])))
```

### TABLA RESÚMEN DE EIGENVALORES DE SDELTA

```
tab=matrix(0, nrow=num.est, ncol=4)
colnames(tab)=c("Componente", "Eigenvalor", "Porcentaje", "%Acumulado")
tab[,1]=1:num.est
tab[,2]=round(val.Sdelta,digits=3)
tab[,3]=round((val.Sdelta/sum(val.Sdelta))*100,digits=1)
tab[,4]=round(cumsum(tab[,3]),digits=1)
```

### GRÁFICA DE LA LADERA

```
plot(tab[,1], tab[,2], main=paste("GRÁFICA DE LA LADERA DE SDELTA"),
xlab="No. de eigenvalor", ylab="Eigenvalor", type="b")
```

### GRÁFICA DE LA IMÁGEN EUCLIDIANA

```
x=seq(length=10, from=0, to=max(pca.Sdelta[, 1]+1))
y=seq(length=10, from=min(pca.Sdelta[, 2]), to=max(pca.Sdelta[, 2]))
plot(x, y, type='n', main="IMÁGEN EUCLIDIANA", xlab="", ylab="")
text(pca.Sdelta[, 1], pca.Sdelta[, 2], cex=1, pos=4, labels=tiempo)
segments(0, 0, pca.Sdelta[, 1], pca.Sdelta[, 2], col='red')
```

### COMPROMISO DE LAS CIUDADES

```
delta.C=diag(peso.est, num.est, num.est)
W =array(0, dim=c(num.obs, num.obs, num.est))
suma =matrix(0, num.obs, num.obs)
traza =sum(sqrt(diag(S)))
```

```

alfa=(1/sqrt(val.Sdelta[1]))*traza*delta.C%%vec.Sdelta[, 1]
for(i in 1:num.est)
{
  W[, , i]=alfa[i, 1]*mpe[, , i]
  suma=suma+W[, , i]
}
WD=sqrt(D)%%suma%%sqrt(D)

```

#### ANÁLISIS DE COMPONENTES PRINCIPALES DE WD

```

comp.WD=eigen(WD, EISPACK=TRUE)
EVE=as.matrix(comp.WD$eigenvectors)
EVA=as.matrix(comp.WD$values)
EVA2=diag(sqrt(abs(1/as.vector(EVA))))
pca.WD=WD%%EVE%%EVA2

```

#### TABLA RESÚMEN DE EIGENVALORES DE WD

```

tab=matrix(0, nrow=num.obs, ncol=4)
colnames(tab)=c("Componente", "Eigenvalor", "Porcentaje", "%Acumulado")
tab[,1]=1:num.obs
tab[,2]=round(EVA,digits=3)
tab[,3]=round((EVA/sum(EVA))*100,digits=1)
tab[,4]=round(cumsum(tab[,3]),digits=1)

```

#### GRÁFICA DE LA LADERA DE WD

```

plot(tab[,1], tab[,2], main=paste("GRÁFICA DE LA LADERA DE WD"),
xlab="No. de eigenvalor", ylab="Eigenvalor", type="b")

```

#### GRÁFICA DEL COMPROMISO DE LAS CIUDADES

```

plot(pca.WD[, 1], pca.WD[, 2], pch=16, type="p", col="red",
main="COMPROMISO DE LAS CIUDADES DEL CORREDOR INDUSTRIAL",
xlab="", ylab="")
text(pca.WD[, 1], pca.WD[, 2], cex=.5, pos=3, labels=ciudad)

```

#### GRÁFICAS DE LAS TRAYECTORIAS

```

coord=array(0, dim=c(num.obs, num.obs, num.est))
X=matrix(0, num.obs, num.est)
Y=matrix(0, num.obs, num.est)
for(i in 1:num.est)
{
  coord[, , i]=mpe[, , i]%%D%%EVE%%EVA2
  X[, i]=coord[, 1, i]
  Y[, i]=coord[, 2, i]
}
trayectorias=function(X, Y, tiempo, n.e, ...)
{
  plot(X[, 1:n.e], Y[, 1:n.e], pch=16, type="p", col="red", xlab="",
ylab="", main=paste("TRAYECTORIAS DE LAS CIUDADES DEL

```

```

CORREDOR INDUSTRIAL\n", "DE ENE A ", tiempo[i], "DE 2006"))
text(X[, 1], Y[, 1], cex=.5, pos=3, labels=ciudad)
if(n.e>1)
{
  for(i in 2:n.e)
    text(X[, i], Y[, i], cex=.5, pos=3, labels=tiempo[i])
  for(i in 1:(n.e-1))
    segments(X[, i], Y[, i], X[, i+1], Y[, i+1], col='red')
}
}

for(i in 1:num.est)
{
  x11()
  trayectorias(X, Y, tiempo, i)
}

```

**CORRELACIONES CON LOS EJES DEL COMPROMISO**

```

coord=array(0, dim=c(num.vars, num.obs, num.est))
X =matrix(0, num.vars, num.est)
Y =matrix(0, num.vars, num.est)
for(i in 1:num.est)
{
  coord[, , i]=(1/sqrt(num.obs))*t(Z[, , i])%%EVE
  X[, i]=coord[, 1, i]
  Y[, i]=coord[, 2, i]
}

```

**TABLA DE CORRELACIONES CON EL COMPROMISO**

```

correlacion =c(" ENE_1 ", " FEB_1 ", " MAR_1 ", " ABR_1 ",
  " MAY_1 ", " JUN_1 ", " JUL_1 ", " AGO_1 ", " SEP_1 ",
  " OCT_1 ", " NOV_1 ", " DIC_1 ", " ENE_2 ", " FEB_2 ",
  " MAR_2 ", " ABR_2 ", " MAY_2 ", " JUN_2 ", " JUL_2 ",
  " AGO_2 ", " SEP_2 ", " OCT_2 ", " NOV_2 ", " DIC_2 ")
tabla2=matrix(0, num.vars, 24)
dimnames(tabla2)=list(vars, correlacion)
for(i in 1:12)
  tabla2[, i]=round(X[, i], digits=2)
for(i in 13:24)
  tabla2[, i]=round(Y[, i-12], digits=2)

```

**GRÁFICA DE CORRELACIONES**

```

XX=seq(-1, 1, 0.5)
YY=seq(-2, 1, 0.75)
colores=c("blue", "brown", "green", "cyan", "red", "black", "purple3",
  "yellow1", "orange", "salmon", "dimgray")
plot(XX, YY, type = "n", main="CORRELACIÓN DE LAS VARIABLES

```



```

CON EL COMPROMISO", xlab="Correlación de Variables con el eje
principal", ylab="Correlación de Variables con el eje secundario")
for(j in 1:num.est)
  for(i in 1:num.vars)
  {
    points(X[i, ], Y[i, ], pch=16, type="p", col = colores[i], cex = 1.0)
    text(X[i, ], Y[i, ], cex=.5, pos=3, labels=tiempo)
    text((X[i, 1]+X[i, 11])/2, (Y[i, 1]+Y[i, 11])/2, cex=1.0,
        pos=3, labels=vars[i])
    for(k in 1:(num.est-1))
      segments(X[i, k], Y[i, k], X[i, k+1], Y[i, k+1], col=colores[i], lwd=2)
  }
op=par(bg="antiquewhite1")
legend(-.95, -1.3, paste(labels=vars[1:12]), col=colores[1:12], lty=1, pt.cex = 1,
      pch =19, lwd=2, ncol=4, cex=0.9, text.col="black", title = "Variables")
par(op)

```

**GRÁFICAS INDIVIDUALES DE CORRELACIONES**

```

trayectorias=function(X,Y,tiempo,n.e,cual,...)
{
  plot(X[,1:n.e], Y[,1:n.e], pch=16, type="n",
      main=paste("CORRELACION DE LA VARIABLE\n",vars[cual]),
      xlab="Correlación de Variables con el eje principal",
      ylab="Correlación de Variables con el eje secundario")
  points(X[cual,1:n.e], Y[cual,1:n.e], pch=16, type="p",
        col = colores[cual],cex = 1.0)
  text(X[cual,1], Y[cual,1], cex=.5, pos=3)
  if(n.e>1)
  {
    for(i in 1:n.e)
      text(X[cual,i], Y[cual,i], cex=.5, pos=3, labels=tiempo[i])
    for(i in 1:(n.e-1))
      segments(X[cual,i], Y[cual,i], X[cual,i+1], Y[cual,i+1],
              col=colores[cual])
  }
}
for(j in 1:num.vars)
  x11()
  for(i in 1:num.est)
    trayectorias(X, Y, tiempo, i,10)
}

```

TESIS TESIS TESIS TESIS TESIS



TESIS TESIS TESIS TESIS TESIS





Julio	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	400.9	4.7	62.0	8.7	3.1	4.2	49.7	30.7	95.3	679.6	995.9
Irapuato	330.6	6.8	165.0	134.8	9.1	95.4	14.3	34.1	99.9	680.0	992.0
Salamanca	102.4	0.7	104.6	38.4	33.0	69.0	6.1	32.4	90.4	830.6	968.1
León	145.8	26.0	177.1	336.0	60.0	353.1	7.1	29.3	91.5	671.6	999.2

Valores Máximos de las Variables en el mes de Julio

Agosto	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	119.7	5.0	182.0	180.6	9.1	189.0	13.1	29.3	99.2	678.9	999.1
Irapuato	367.3	34.5	358.1	342.2	35.2	323.7	5.7	28.0	99.9	626.6	808.5
Salamanca	107.7	1.2	433.5	48.4	26.0	79.0	4.9	28.9	91.3	826.5	968.1
León	183.8	7.5	123.6	477.8	64.2	491.8	6.5	27.0	91.8	670.0	986.6

Valores Máximos de las Variables en el mes de Agosto

Septiembre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	135.7	1.0	3.0	101.1	108.8	152.9	11.9	28.6	99.5	679.2	998.5
Irapuato	100.2	5.2	104.8	136.5	34.5	152.4	5.3	28.2	99.9	627.2	799.5
Salamanca	135.0	13.2	492.9	72.9	39.9	96.4	5.5	30.2	94.6	830.3	991.6
León	136.4	6.1	93.8	125.7	64.2	123.6	5.8	27.9	91.1	670.0	999.4

Valores Máximos de las Variables en el mes de Septiembre

Octubre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	333.8	1.8	3.0	27.1	3.1	27.0	10.7	29.9	99.5	680.3	998.1
Irapuato	118.8	6.5	127.8	221.7	51.2	258.4	5.8	29.2	99.9	629.0	768.0
Salamanca	353.2	13.0	457.5	95.0	51.0	160.0	5.8	29.7	95.2	833.1	984.1
León	100.2	6.1	136.4	125.7	64.2	123.6	7.0	27.5	91.4	670.8	999.2

Valores Máximos de las Variables en el mes de Octubre

Noviembre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	208.4	4.6	102.0	174.9	482.9	489.6	12.6	28.8	99.6	681.1	907.6
Irapuato	240.0	0.3	200.2	397.1	93.9	449.7	18.1	28.2	92.7	705.1	667.2
Salamanca	305.0	18.2	441.0	171.0	67.0	232.0	64.5	28.4	93.8	834.5	856.1
León	249.8	18.7	201.0	386.0	18.8	389.5	6.9	26.5	85.2	673.5	996.4

Valores Máximos de las Variables en el mes de Noviembre

Diciembre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	116.0	2.8	58.0	83.9	41.8	97.7	16.1	26.1	99.6	680.9	863.4
Irapuato	153.6	8.7	449.7	350.3	99.2	422.4	16.6	26.4	99.9	703.9	788.1
Salamanca	100.5	7.9	422.0	205.0	60.0	245.0	8.0	26.4	83.8	834.3	761.9
León	126.4	5.2	89.6	5.7	0.8	5.4	6.7	30.3	86.4	672.3	989.8

Valores Máximos de las Variables en el mes de Diciembre

## TABLAS DE DATOS ESTANDARIZADOS

Marzo	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	-0.30	0.09	-0.69	-0.99	-0.82	-1.07	1.02	-0.55	1.68	-0.54	0.73
Irapuato	-0.28	0.41	0.08	0.93	1.60	0.88	-0.86	-0.63	-0.16	-0.52	0.65
Salamanca	-1.06	1.11	1.60	1.07	0.10	1.11	0.98	-0.55	-0.67	1.73	-1.71
León	1.64	-1.61	-0.98	-1.01	-0.88	-0.92	-1.13	1.73	-0.84	-0.67	0.33

Datos Estandarizados del mes de Marzo

Abril	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	-0.62	-0.73	-1.29	-0.74	-0.85	-0.94	1.70	-0.61	0.15	-0.54	-0.11
Irapuato	1.60	1.70	0.92	1.68	1.62	1.59	-0.29	0.61	1.20	-0.52	1.30
Salamanca	-1.02	-0.73	1.03	-0.18	0.05	0.10	-0.76	-1.28	0.22	1.73	-1.49
León	0.04	-0.24	-0.66	-0.76	-0.82	-0.75	-0.66	1.28	-1.58	-0.67	0.29

Datos Estandarizados del mes de Abril

Mayo	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	1.37	1.00	-1.18	1.03	1.69	1.44	1.71	-0.21	0.52	-0.54	1.35
Irapuato	0.32	-0.14	-0.76	0.78	-0.79	-0.20	-0.29	-0.38	1.36	-0.52	-0.35
Salamanca	-1.39	0.71	1.26	-0.32	-0.21	0.12	-0.71	-1.05	-0.72	1.73	-1.38
León	-0.30	-1.57	0.68	-1.49	-0.69	-1.36	-0.70	1.64	-1.17	-0.67	0.39

Datos Estandarizados del mes de Mayo

Junio	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	-1.01	-0.61	-1.32	-1.12	-0.72	-1.10	1.71	-0.79	0.05	-0.54	0.71
Irapuato	1.57	-0.27	-0.56	0.46	-0.64	-0.11	-0.29	0.31	1.62	-0.52	0.19
Salamanca	-0.69	-0.82	1.24	-0.75	-0.36	-0.40	-0.72	-1.02	-0.87	1.73	-1.69
León	0.13	1.70	0.64	1.40	1.72	1.62	-0.70	1.50	-0.79	-0.67	0.78

Datos Estandarizados del mes de Junio

Julio	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	1.25	-0.50	-1.40	-0.94	-1.03	-0.95	1.70	-0.51	0.28	-0.54	0.58
Irapuato	0.69	-0.28	0.81	0.04	-0.77	-0.26	-0.28	1.37	1.51	-0.53	0.26
Salamanca	-1.15	-0.91	-0.48	-0.71	0.30	-0.46	-0.74	0.43	-1.04	1.73	-1.69
León	-0.80	1.69	1.07	1.61	1.50	1.68	-0.68	-1.29	-0.75	-0.66	0.85

Datos Estandarizados del mes de Julio

Agosto	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	-0.72	-0.54	-0.73	-0.50	-1.23	-0.53	1.71	1.13	0.91	-0.29	0.76
Irapuato	1.66	1.71	0.66	0.49	0.08	0.34	-0.57	-0.34	1.08	-0.98	-1.71
Salamanca	-0.84	-0.82	1.26	-1.32	-0.38	-1.24	-0.81	0.68	-1.06	1.67	0.36
León	-0.10	-0.35	-1.20	1.33	1.53	1.43	-0.32	-1.47	-0.93	-0.40	0.60

Datos Estandarizados del mes de Agosto

Septiembre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	0.58	-1.23	-0.91	-0.32	1.60	0.92	1.73	-0.14	0.88	-0.29	0.60
Irapuato	-1.73	-0.27	-0.37	1.12	-0.93	0.90	-0.66	-0.59	0.99	-0.97	-1.73
Salamanca	0.53	1.56	1.69	-1.48	-0.75	-1.49	-0.59	1.66	-0.46	1.67	0.52
León	0.62	-0.06	-0.42	0.68	0.08	-0.33	-0.48	-0.93	-1.42	-0.41	0.61

Datos Estandarizados del mes de Septiembre

Octubre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NOX	WS	TEMP	HR	PBAR	RADSOL
Celaya	0.91	-1.26	-1.06	-1.29	-1.69	-1.39	1.68	0.87	0.86	-0.30	0.62
Irapuato	-0.92	-0.09	-0.32	1.49	0.38	1.40	-0.76	0.13	0.98	-0.96	-1.73
Salamanca	1.08	1.54	1.64	-0.32	0.37	0.21	-0.76	0.66	-0.37	1.68	0.48
León	-1.08	-0.19	-0.27	0.12	0.94	-0.23	-0.16	-1.67	-1.47	-0.42	0.63

Datos Estandarizados del mes de Octubre

Noviembre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
Celaya	-1.22	-0.72	-1.07	-0.98	1.71	1.01	-0.57	0.94	1.32	-0.65	0.42
Irapuato	-0.31	-1.25	-0.29	1.05	-0.39	0.61	-0.32	0.26	-0.02	-0.28	-1.57
Salamanca	1.56	0.95	1.64	-1.02	-0.53	-1.61	1.71	0.48	0.19	1.70	-0.01
León	-0.03	1.01	-0.28	0.95	-0.79	-0.01	-0.82	-1.68	-1.49	-0.77	1.16

Datos Estandarizados del mes de Noviembre

Diciembre	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
Celaya	-0.42	-1.44	-1.08	-0.59	-0.24	-0.60	0.94	-0.69	0.97	-0.64	0.14
Irapuato	1.52	1.10	1.07	1.45	1.38	1.46	1.05	-0.52	1.01	-0.29	-0.71
Salamanca	-1.22	0.75	0.92	0.34	0.27	0.33	-0.85	-0.52	-1.17	1.70	-1.00
León	0.12	-0.41	-0.91	-1.19	-1.40	-1.19	-1.14	1.73	-0.82	-0.77	1.57

Datos Estandarizados del mes de Diciembre

### TABLAS DE CORRELACIÓN POR MES

Marzo	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O <sub>3</sub>	1.00	-0.99	-0.78	-0.69	-0.44	-0.66	-0.74	0.94	-0.28	-0.66	0.49
CO	-0.99	1.00	0.83	0.78	0.53	0.75	0.66	-0.93	0.17	0.68	-0.53
SO <sub>2</sub>	-0.78	0.83	1.00	0.86	0.43	0.87	0.47	-0.56	-0.36	0.94	-0.88
NO	-0.69	0.78	0.86	1.00	0.82	1.00	0.10	-0.59	-0.42	0.65	-0.57
NO <sub>2</sub>	-0.44	0.53	0.43	0.82	1.00	0.80	-0.28	-0.53	-0.24	0.09	0.00
NO <sub>x</sub>	-0.66	0.75	0.87	1.00	0.80	1.00	0.07	-0.54	-0.48	0.67	-0.60
WS	-0.74	0.66	0.47	0.10	-0.28	0.07	1.00	-0.63	0.53	0.59	-0.47
TEMP	0.94	-0.93	-0.56	-0.59	-0.53	-0.54	-0.63	1.00	-0.48	-0.37	0.18
HR	-0.28	0.17	-0.36	-0.42	-0.24	-0.48	0.53	-0.48	1.00	-0.35	0.50
PBAR	-0.66	0.68	0.94	0.65	0.09	0.67	0.59	-0.37	-0.35	1.00	-0.98
RADSOL	0.49	-0.53	-0.88	-0.57	0.00	-0.60	-0.47	0.18	0.50	-0.98	1.00

Matriz de Correlaciones de las Variables en Marzo

Abril	O <sub>3</sub>	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O <sub>3</sub>	1.00	0.98	0.30	0.83	0.76	0.75	-0.19	0.68	0.39	-0.57	0.92
CO	0.98	1.00	0.48	0.93	0.88	0.87	-0.25	0.52	0.54	-0.40	0.82
SO <sub>2</sub>	0.30	0.48	1.00	0.71	0.79	0.82	-0.70	-0.20	0.54	0.61	-0.10
NO	0.83	0.93	0.71	1.00	0.99	0.98	-0.28	0.19	0.77	-0.07	0.58
NO <sub>2</sub>	0.76	0.88	0.79	0.99	1.00	1.00	-0.35	0.10	0.78	0.06	0.47
NO <sub>x</sub>	0.75	0.87	0.82	0.98	1.00	1.00	-0.41	0.11	0.74	0.09	0.45
WS	-0.19	-0.25	-0.70	-0.28	-0.35	-0.41	1.00	-0.27	0.20	-0.41	0.10
TEMP	0.68	0.52	-0.20	0.19	0.10	0.11	-0.27	1.00	-0.41	-0.76	0.78
HR	0.39	0.54	0.54	0.77	0.78	0.74	0.20	-0.41	1.00	0.18	0.19
PBAR	-0.57	-0.40	0.61	-0.07	0.06	0.09	-0.41	-0.76	0.18	1.00	-0.85
RADSOL	0.92	0.82	-0.10	0.58	0.47	0.45	0.10	0.78	0.19	-0.85	1.00

Matriz de Correlaciones de las Variables en Abril

Mayo	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.20	-0.95	0.64	0.64	0.54	0.86	0.14	0.63	-0.77	0.88
CO	0.20	1.00	-0.31	0.76	0.69	0.93	0.59	-0.87	0.42	0.45	-0.05
SO <sub>2</sub>	-0.95	-0.31	1.00	-0.80	-0.53	-0.58	-0.79	0.08	-0.84	0.69	-0.70
NO	0.64	0.76	-0.80	1.00	0.56	0.83	0.70	-0.66	0.89	-0.12	0.24
NO <sub>2</sub>	0.64	0.69	-0.53	0.56	1.00	0.88	0.94	-0.24	0.19	-0.10	0.64
NO <sub>x</sub>	0.54	0.93	-0.58	0.83	0.88	1.00	0.85	-0.65	0.50	0.11	0.33
WS	0.86	0.59	-0.79	0.70	0.94	0.85	1.00	-0.16	0.46	-0.38	0.78
TEMP	0.14	-0.87	0.08	-0.66	-0.24	-0.65	-0.16	1.00	-0.45	-0.65	0.49
HR	0.63	0.42	-0.84	0.89	0.19	0.50	0.46	-0.45	1.00	-0.36	0.19
PBAR	-0.77	0.45	0.69	-0.12	-0.10	0.11	-0.38	-0.65	-0.36	1.00	-0.80
RADSOL	0.88	-0.05	-0.70	0.24	0.64	0.33	0.78	0.49	0.19	-0.80	1.00

Matriz de Correlaciones de las Variables en Mayo

Junio	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.25	-0.08	0.64	0.05	0.36	-0.45	0.55	0.75	-0.39	0.21
CO	0.25	1.00	0.26	0.89	0.95	0.94	-0.39	0.94	-0.27	-0.52	0.56
SO <sub>2</sub>	-0.08	0.26	1.00	0.30	0.49	0.52	-0.86	0.14	-0.64	0.68	-0.66
NO	0.64	0.89	0.30	1.00	0.80	0.94	-0.62	0.97	0.06	-0.47	0.41
NO <sub>2</sub>	0.05	0.95	0.49	0.80	1.00	0.95	-0.49	0.83	-0.53	-0.26	0.33
NO <sub>x</sub>	0.36	0.94	0.52	0.94	0.95	1.00	-0.67	0.91	-0.29	-0.28	0.28
WS	-0.45	-0.39	-0.86	-0.62	-0.49	-0.67	1.00	-0.44	0.20	-0.39	0.46
TEMP	0.55	0.94	0.14	0.97	0.83	0.91	-0.44	1.00	0.04	-0.63	0.60
HR	0.75	-0.27	-0.64	0.06	-0.53	-0.29	0.20	0.04	1.00	-0.46	0.30
PBAR	-0.39	-0.52	0.68	-0.47	-0.26	-0.28	-0.39	-0.63	-0.46	1.00	-0.98
RADSOL	0.21	0.56	-0.66	0.41	0.33	0.28	0.46	0.60	0.30	-0.98	1.00

Matriz de Correlaciones de las Variables en Junio

Julio	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	-0.28	-0.37	-0.41	-0.84	-0.55	0.83	0.21	0.79	-0.63	0.54
CO	-0.28	1.00	0.68	0.96	0.75	0.95	-0.31	-0.67	-0.22	-0.57	0.65
SO <sub>2</sub>	-0.37	0.68	1.00	0.86	0.57	0.78	-0.75	0.06	0.14	-0.31	0.28
NO	-0.41	0.96	0.86	1.00	0.79	0.98	-0.55	-0.46	-0.16	-0.45	0.51
NO <sub>2</sub>	-0.84	0.75	0.57	0.79	1.00	0.89	-0.70	-0.58	-0.72	0.12	-0.01
NO <sub>x</sub>	-0.55	0.95	0.78	0.98	0.89	1.00	-0.59	-0.56	-0.36	-0.31	0.40
WS	0.83	-0.31	-0.75	-0.55	-0.70	-0.59	1.00	-0.17	0.33	-0.40	0.40
TEMP	0.21	-0.67	0.06	-0.46	-0.58	-0.56	-0.17	1.00	0.61	0.28	-0.44
HR	0.79	-0.22	0.14	-0.16	-0.72	-0.36	0.33	0.61	1.00	-0.57	0.42
PBAR	-0.63	-0.57	-0.31	-0.45	0.12	-0.31	-0.40	0.28	-0.57	1.00	-0.99
RADSOL	0.54	0.65	0.28	0.51	-0.01	0.40	0.40	-0.44	0.42	-0.99	1.00

Matriz de Correlaciones de las Variables en Julio

Agosto	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.99	0.18	0.54	0.29	0.46	-0.37	-0.45	0.53	-0.70	-0.94
CO	0.99	1.00	0.23	0.43	0.14	0.35	-0.28	-0.31	0.64	-0.69	-0.96
SO <sub>2</sub>	0.18	0.23	1.00	-0.64	-0.34	-0.67	-0.57	0.39	-0.04	0.54	-0.49
NO	0.54	0.43	-0.64	1.00	0.80	1.00	-0.12	-0.89	0.06	-0.77	-0.23
NO <sub>2</sub>	0.29	0.14	-0.34	0.80	1.00	0.84	-0.58	-0.98	-0.51	-0.25	-0.07
NO <sub>x</sub>	0.46	0.35	-0.67	1.00	0.84	1.00	-0.14	-0.91	-0.03	-0.71	-0.14
WS	-0.37	-0.28	-0.57	-0.12	-0.58	-0.14	1.00	0.51	0.52	-0.29	0.45
TEMP	-0.45	-0.31	0.39	-0.89	-0.98	-0.91	0.51	1.00	0.33	0.43	0.20
HR	0.53	0.64	-0.04	0.06	-0.51	-0.03	0.52	0.33	1.00	-0.68	-0.53
PBAR	-0.70	-0.69	0.54	-0.77	-0.25	-0.71	-0.29	0.43	-0.68	1.00	0.45
RADSOL	-0.94	-0.96	-0.49	-0.23	-0.07	-0.14	0.45	0.20	-0.53	0.45	1.00

Matriz de Correlaciones de las Variables en Agosto



Septiembre	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.14	0.19	-0.62	0.55	-0.51	0.38	0.31	-0.58	0.54	1.00
CO	0.14	1.00	0.97	-0.56	-0.72	-0.92	-0.71	0.74	-0.49	0.81	0.12
SO <sub>2</sub>	0.19	0.97	1.00	-0.73	-0.60	-0.89	-0.53	0.89	-0.34	0.91	0.18
NO	-0.62	-0.56	-0.73	1.00	-0.10	0.67	-0.19	-0.93	0.14	-0.94	-0.62
NO <sub>2</sub>	0.55	-0.72	-0.60	-0.10	1.00	0.43	0.95	-0.25	0.18	-0.21	0.56
NO <sub>x</sub>	-0.51	-0.92	-0.89	0.67	0.43	1.00	0.51	-0.71	0.72	-0.88	-0.50
WS	0.38	-0.71	-0.53	-0.19	0.95	0.51	1.00	-0.10	0.46	-0.16	0.40
TEMP	0.31	0.74	0.89	-0.93	-0.25	-0.71	-0.10	1.00	-0.04	0.95	0.31
HR	-0.58	-0.49	-0.34	0.14	0.18	0.72	0.46	-0.04	1.00	-0.35	-0.57
PBAR	0.54	0.81	0.91	-0.94	-0.21	-0.88	-0.16	0.95	-0.35	1.00	0.53
RADSOL	1.00	0.12	0.18	-0.62	0.56	-0.50	0.40	0.31	-0.57	0.53	1.00

Matriz de Correlaciones de las Variables en Septiembre

Octubre	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.20	0.35	-0.75	-0.62	-0.52	0.40	0.80	0.27	0.72	0.50
CO	0.20	1.00	0.99	0.25	0.62	0.50	-0.80	0.05	-0.37	0.78	0.00
SO <sub>2</sub>	0.35	0.99	1.00	0.08	0.51	0.36	-0.69	0.14	-0.36	0.87	0.13
NO	-0.75	0.25	0.08	1.00	0.68	0.95	-0.77	-0.33	0.07	-0.41	-0.86
NO <sub>2</sub>	-0.62	0.62	0.51	0.68	1.00	0.69	-0.89	-0.68	-0.65	0.09	-0.23
NO <sub>x</sub>	-0.52	0.50	0.36	0.95	0.69	1.00	-0.88	-0.13	0.11	-0.12	-0.83
WS	0.40	-0.80	-0.69	-0.77	-0.89	-0.88	1.00	0.28	0.31	-0.24	0.47
TEMP	0.80	0.05	0.14	-0.33	-0.68	-0.13	0.28	1.00	0.77	0.36	-0.11
HR	0.27	-0.37	-0.36	0.07	-0.65	0.11	0.31	0.77	1.00	-0.30	-0.57
PBAR	0.72	0.78	0.87	-0.41	0.09	-0.12	-0.24	0.36	-0.30	1.00	0.50
RADSOL	0.50	0.00	0.13	-0.86	-0.23	-0.83	0.47	-0.11	-0.57	0.50	1.00

Matriz de Correlaciones de las Variables en Octubre

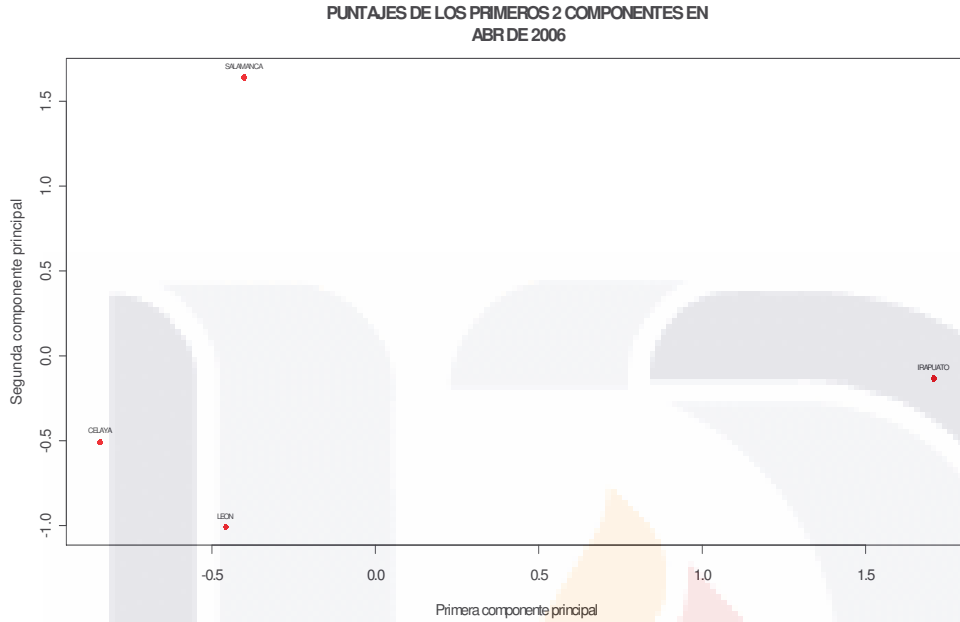
Noviembre	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.68	0.99	-0.19	-0.69	-0.98	0.87	-0.11	-0.32	0.89	-0.02
CO	0.68	1.00	0.60	-0.15	-0.51	-0.76	0.40	-0.56	-0.56	0.42	0.71
SO <sub>2</sub>	0.99	0.60	1.00	-0.30	-0.59	-0.98	0.93	0.05	-0.17	0.95	-0.08
NO	-0.19	-0.15	-0.30	1.00	-0.57	0.32	-0.57	-0.68	-0.73	-0.53	-0.24
NO <sub>2</sub>	-0.69	-0.51	-0.59	-0.57	1.00	0.59	-0.28	0.65	0.84	-0.33	0.10
NO <sub>x</sub>	-0.98	-0.76	-0.98	0.32	0.59	1.00	-0.88	0.08	0.26	-0.89	-0.13
WS	0.87	0.40	0.93	-0.57	-0.28	-0.88	1.00	0.39	0.20	1.00	-0.17
TEMP	-0.11	-0.56	0.05	-0.68	0.65	0.08	0.39	1.00	0.96	0.36	-0.49
HR	-0.32	-0.56	-0.17	-0.73	0.84	0.26	0.20	0.96	1.00	0.15	-0.28
PBAR	0.89	0.42	0.95	-0.53	-0.33	-0.89	1.00	0.36	0.15	1.00	-0.18
RADSOL	-0.02	0.71	-0.08	-0.24	0.10	-0.13	-0.17	-0.49	-0.28	-0.18	1.00

Matriz de Correlaciones de las Variables en Noviembre

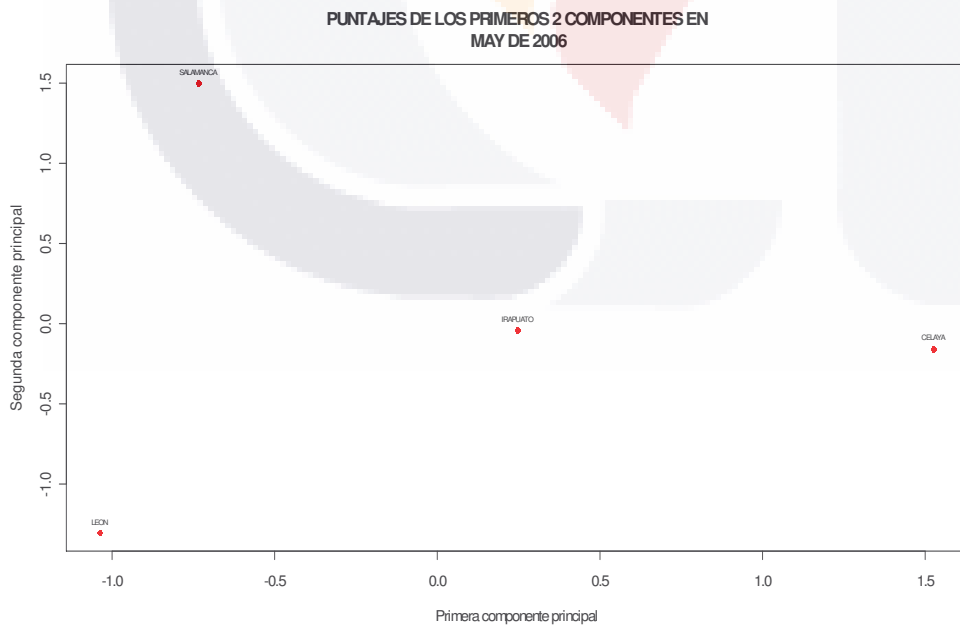
Diciembre	O3	CO	SO <sub>2</sub>	NO	NO <sub>2</sub>	NO <sub>x</sub>	WS	TEMP	HR	PBAR	RADSOL
O3	1.00	0.33	0.21	0.48	0.43	0.48	0.53	0.08	0.62	-0.59	0.07
CO	0.33	1.00	0.95	0.80	0.66	0.80	-0.09	-0.17	-0.21	0.55	-0.59
SO <sub>2</sub>	0.21	0.95	1.00	0.90	0.82	0.90	0.09	-0.46	-0.07	0.66	-0.82
NO	0.48	0.80	0.90	1.00	0.98	1.00	0.51	-0.64	0.37	0.36	-0.83
NO <sub>2</sub>	0.43	0.66	0.82	0.98	1.00	0.98	0.65	-0.78	0.50	0.33	-0.87
NO <sub>x</sub>	0.48	0.80	0.90	1.00	0.98	1.00	0.51	-0.64	0.37	0.36	-0.83
WS	0.53	-0.09	0.09	0.51	0.65	0.51	1.00	-0.68	0.97	-0.37	-0.39
TEMP	0.08	-0.17	-0.46	-0.64	-0.78	-0.64	-0.68	1.00	-0.50	-0.41	0.88
HR	0.62	-0.21	-0.07	0.37	0.50	0.37	0.97	-0.50	1.00	-0.57	-0.17
PBAR	-0.59	0.55	0.66	0.36	0.33	0.36	-0.37	-0.41	-0.57	1.00	-0.70
RADSOL	0.07	-0.59	-0.82	-0.83	-0.87	-0.83	-0.39	0.88	-0.17	-0.70	1.00

Matriz de Correlaciones de las Variables en Diciembre

GRÁFICAS DE LOS PUNTAJES DE LOS DOS PRIMEROS COMPONENTES

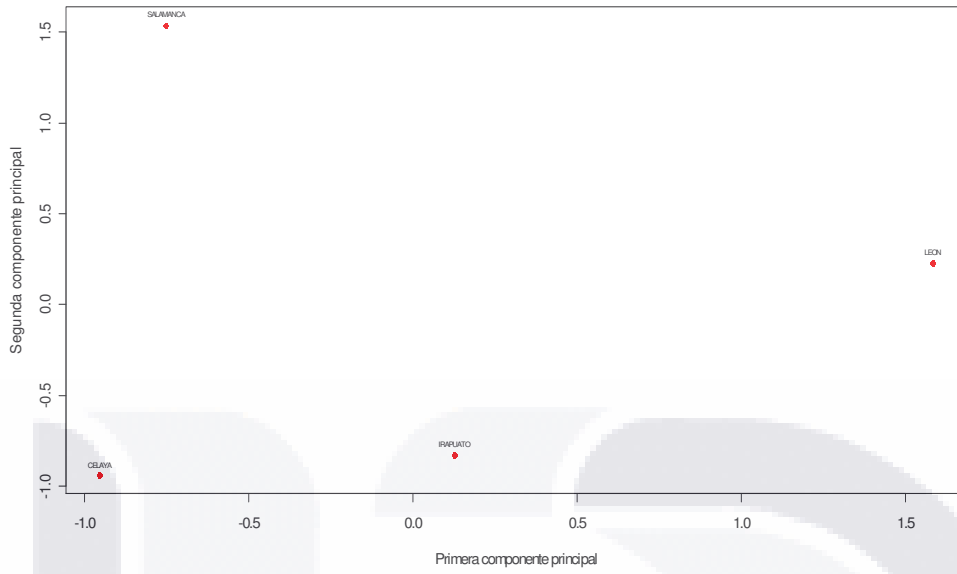


Puntajes de las Ciudades de Guanajuato en Abril



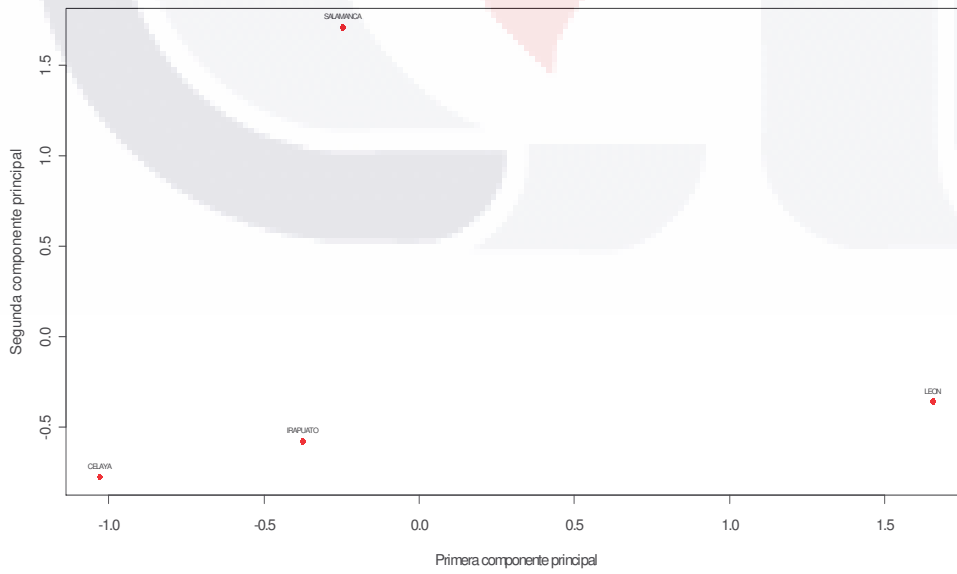
Puntajes de las Ciudades de Guanajuato en Mayo

PUNTAJES DE LOS PRIMEROS 2 COMPONENTES EN JUN DE 2006

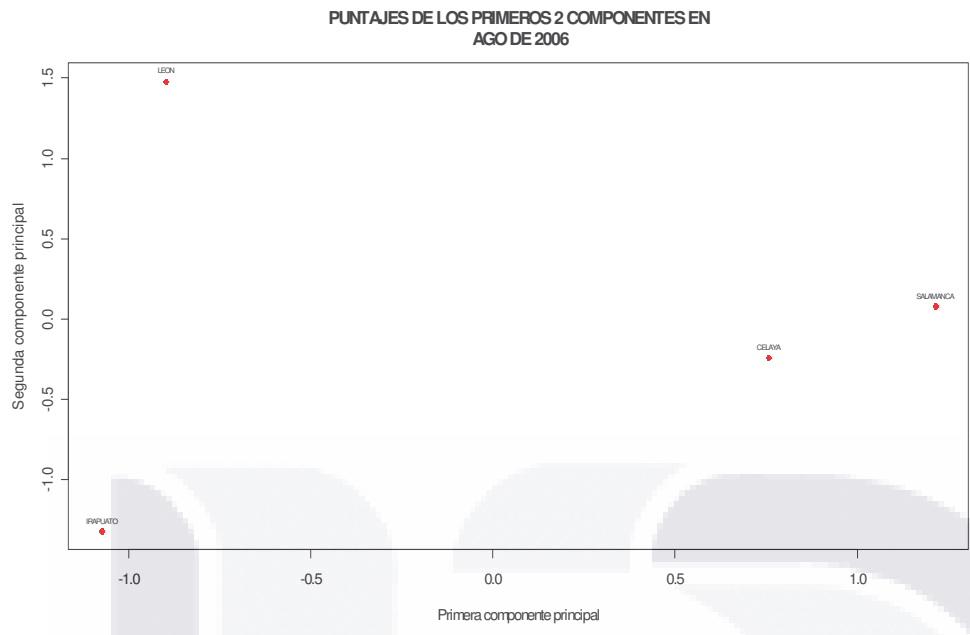


Puntajes de las Ciudades de Guanajuato en Junio

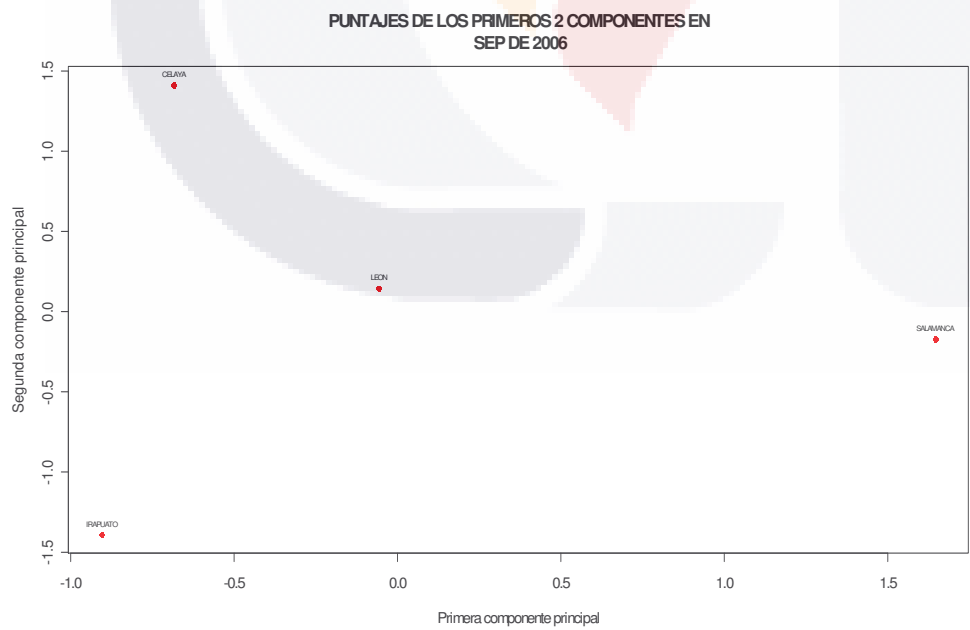
PUNTAJES DE LOS PRIMEROS 2 COMPONENTES EN JUL DE 2006



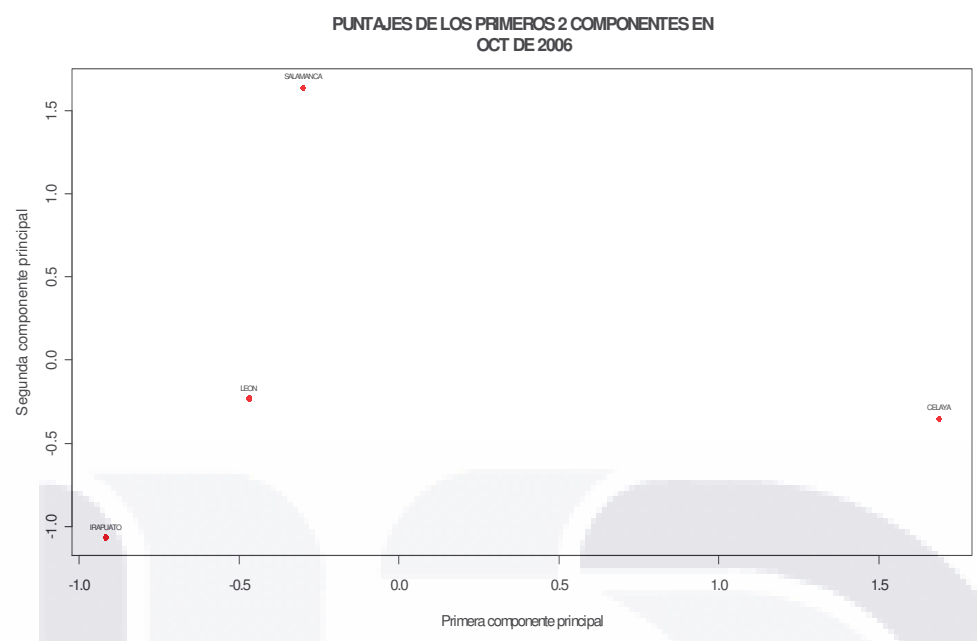
Puntajes de las Ciudades de Guanajuato en Julio



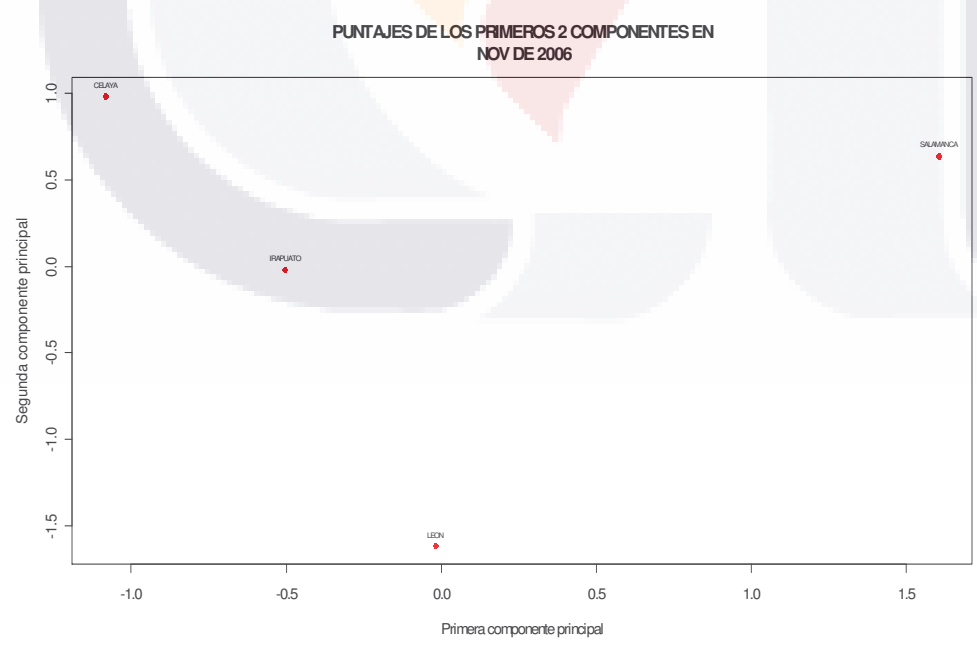
Puntajes de las Ciudades de Guanajuato en Agosto



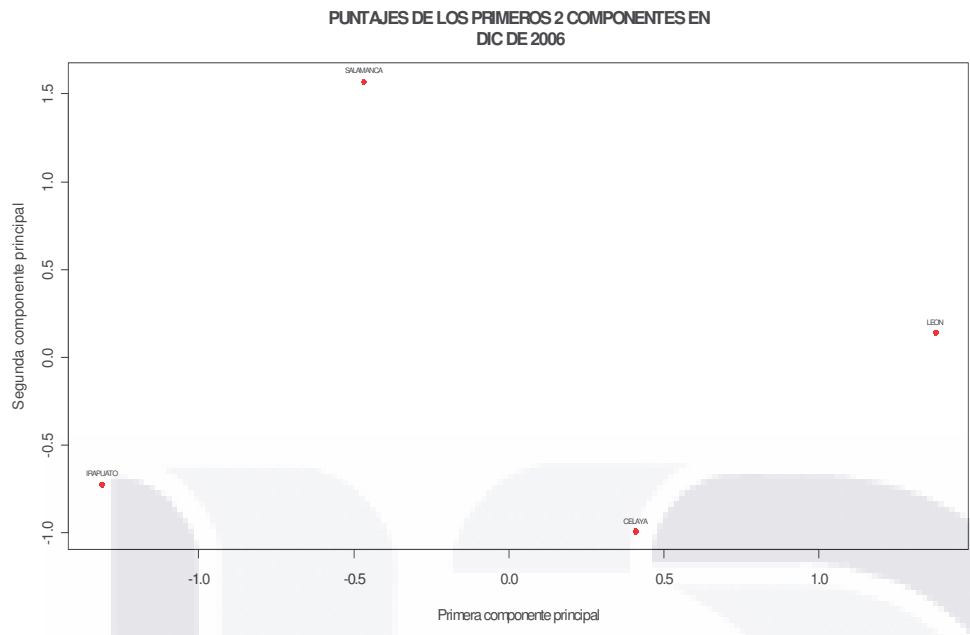
Puntajes de las Ciudades de Guanajuato en Septiembre



Puntajes de las Ciudades de Guanajuato en Octubre

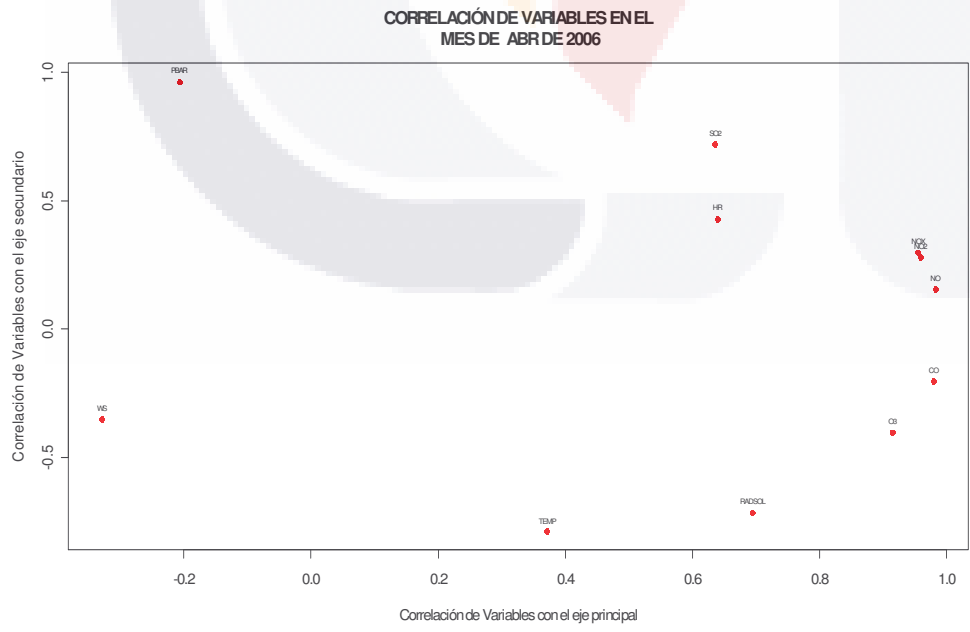


Puntajes de las Ciudades de Guanajuato en Noviembre



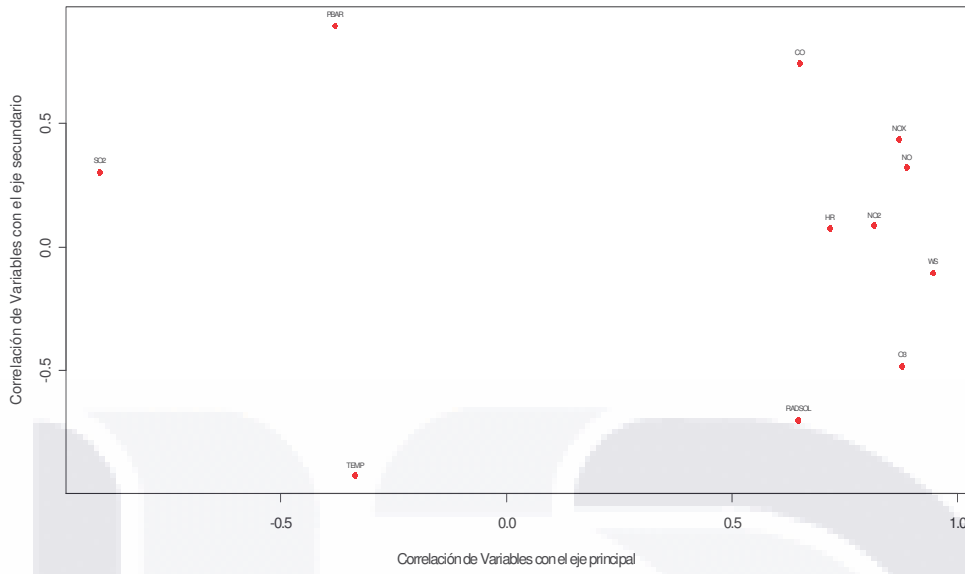
Puntajes de las Ciudades de Guanajuato en Diciembre

### GRÁFICAS DE CORRELACIÓN DE VARIABLES POR MES



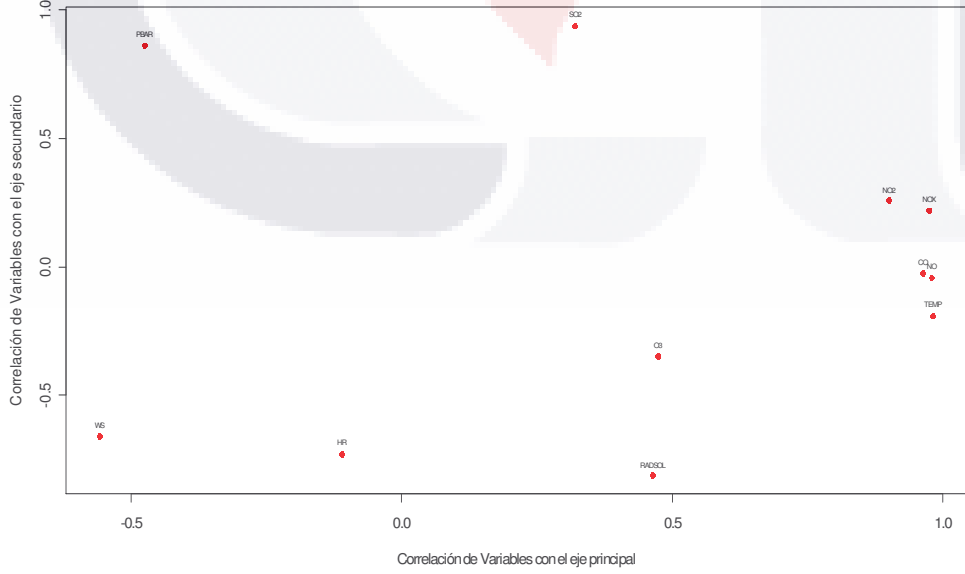
Correlaciones entre Variables en Abril

CORRELACIÓN DE VARIABLES EN EL MES DE MAY DE 2006

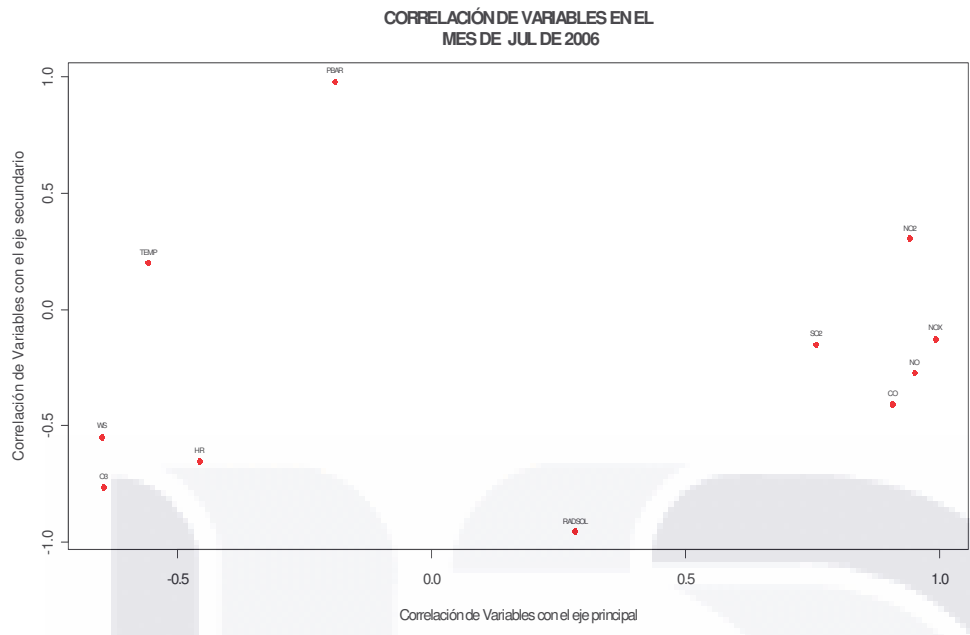


Correlaciones entre Variables en Mayo

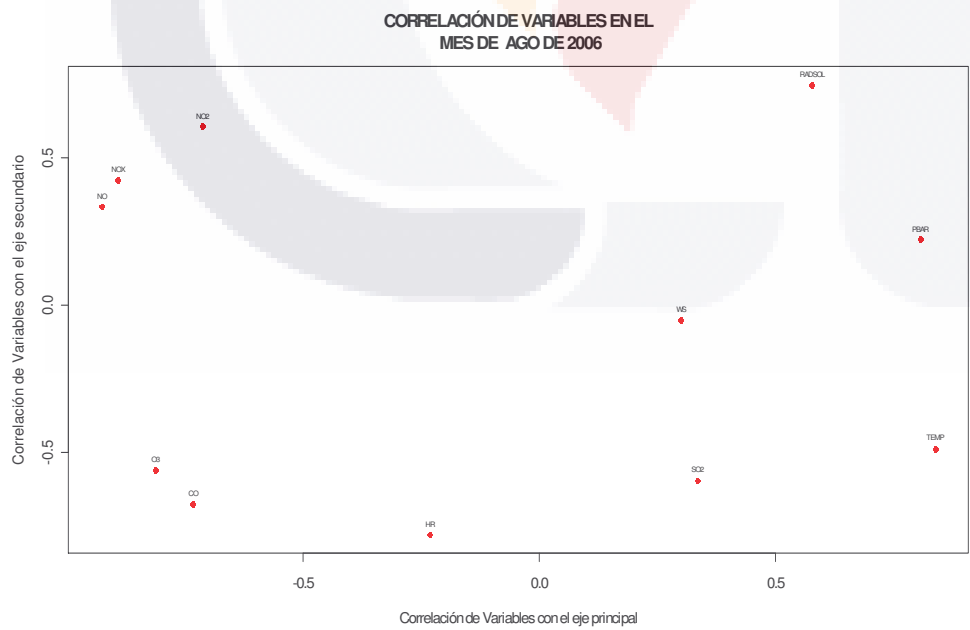
CORRELACIÓN DE VARIABLES EN EL MES DE JUN DE 2006



Correlaciones entre Variables en Junio



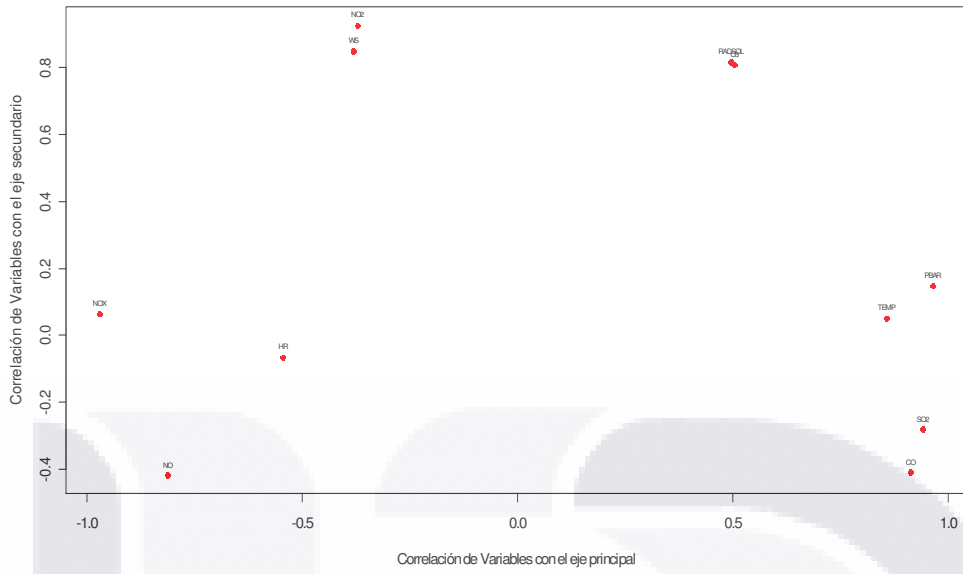
Correlaciones entre Variables en Julio



Correlaciones entre Variables en Agosto

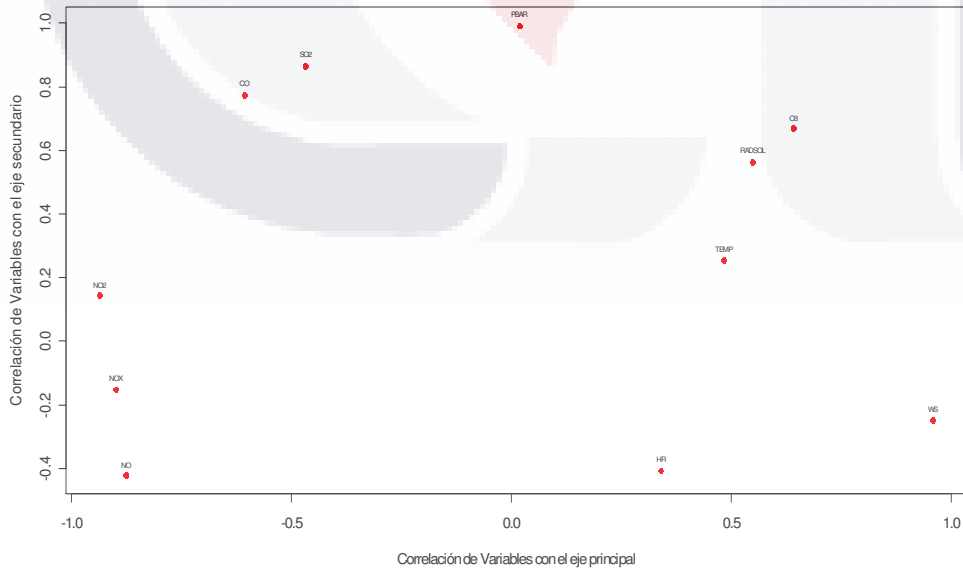


CORRELACIÓN DE VARIABLES EN EL MES DE SEP DE 2006

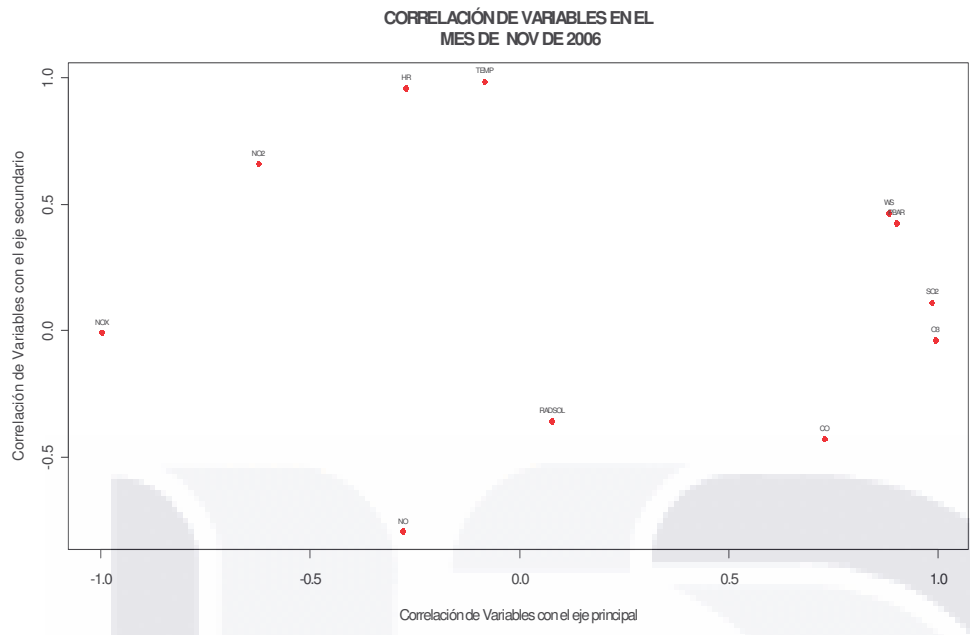


Correlaciones entre Variables en Septiembre

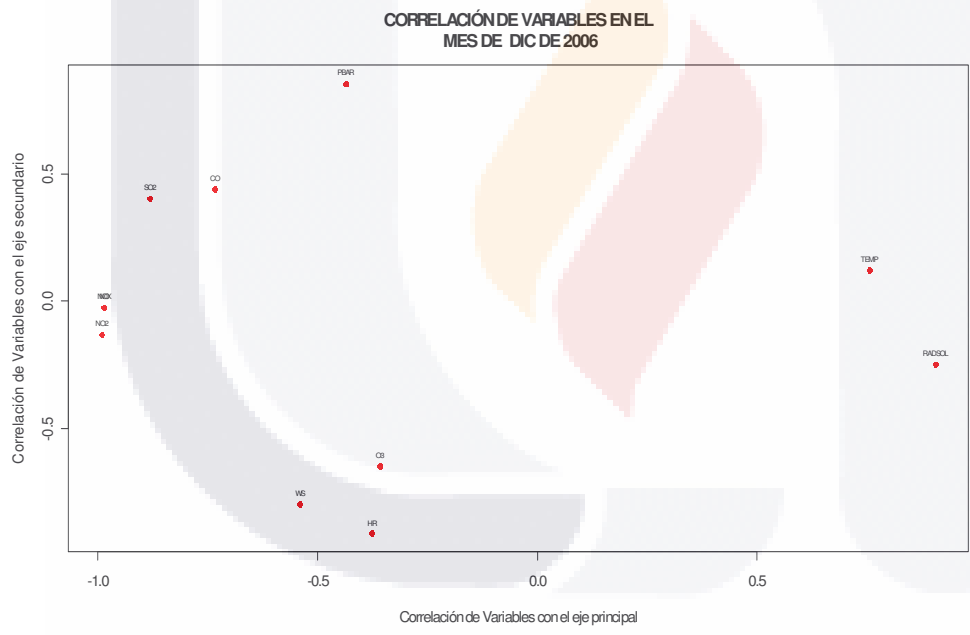
CORRELACIÓN DE VARIABLES EN EL MES DE OCT DE 2006



Correlaciones entre Variables en Octubre



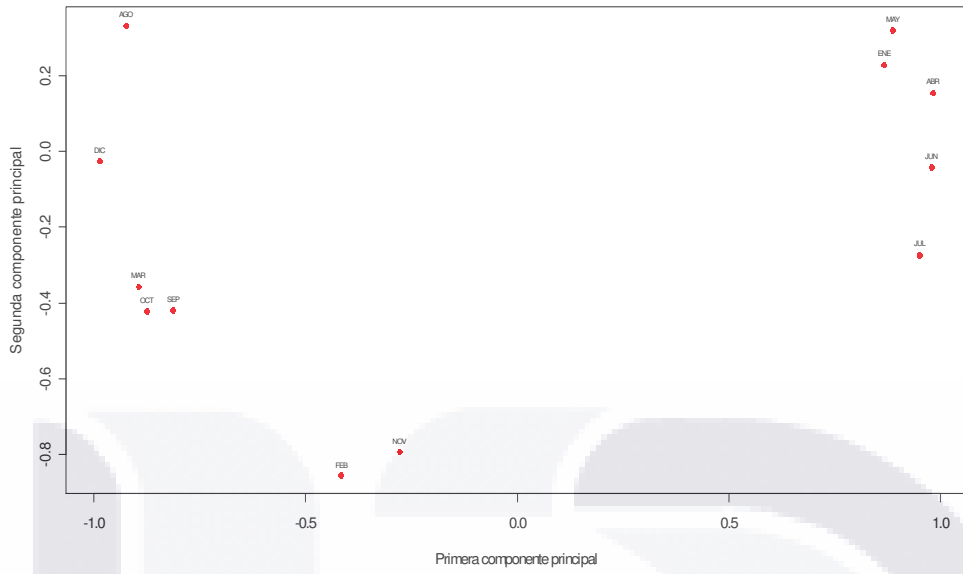
Correlaciones entre Variables en Noviembre



Correlaciones entre Variables en Diciembre

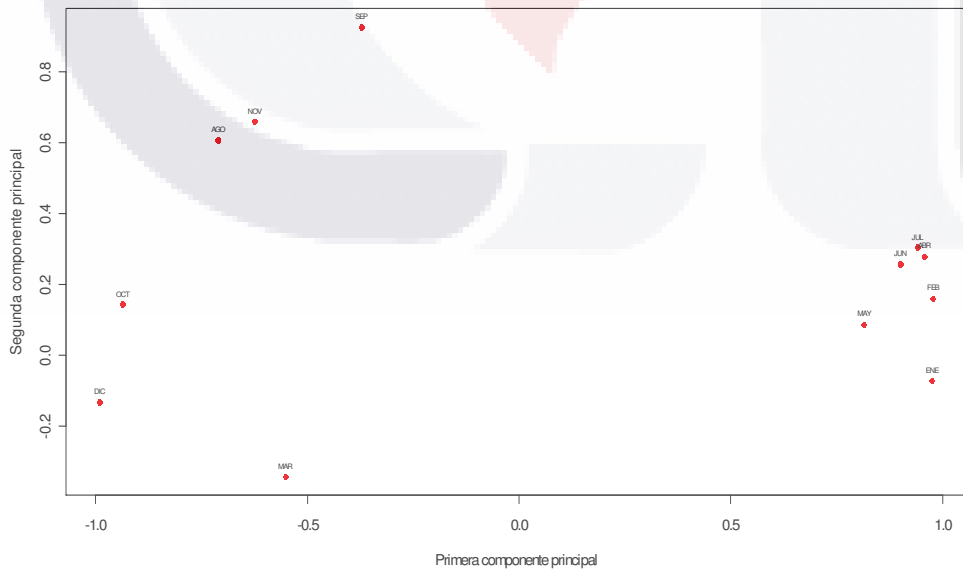
**GRÁFICAS DE CORRELACIONES MENSUALES POR VARIABLE**

CORRELACIONES DE LA VARIABLE NO



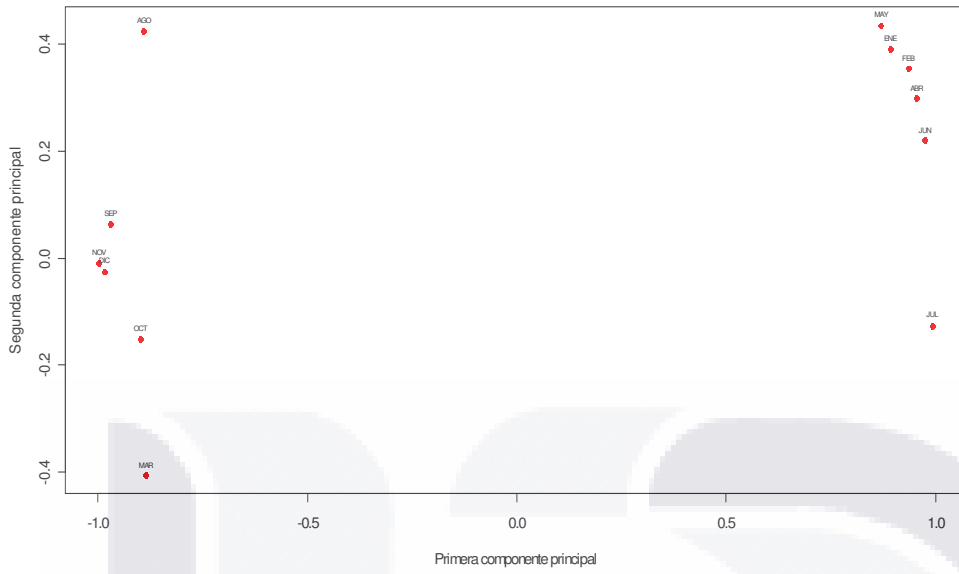
Correlaciones del Monóxido de Nitrógeno a través de los meses

CORRELACIONES DE LA VARIABLE NO2



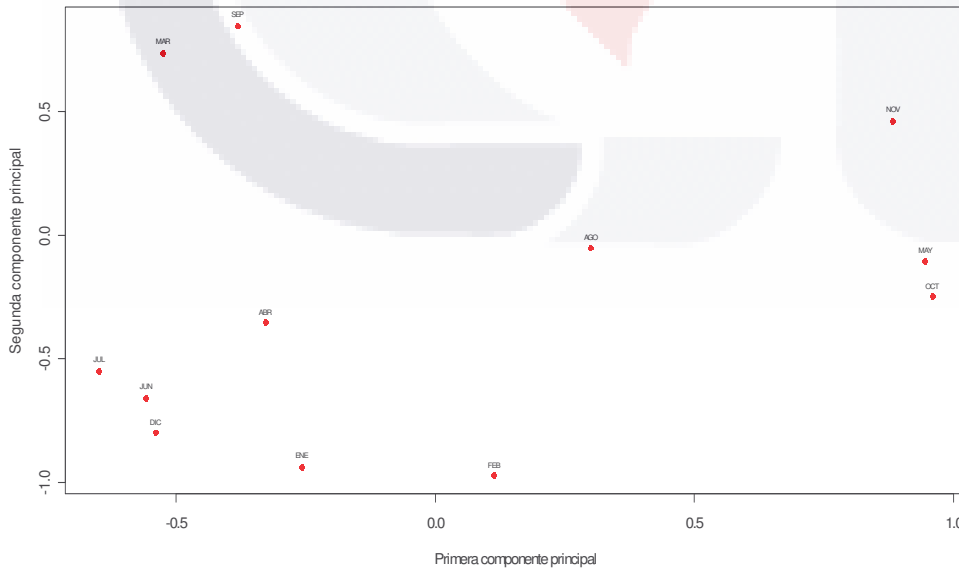
Correlaciones del Dióxido de Nitrógeno a través de los meses

CORRELACIONES DE LA VARIABLE NOX

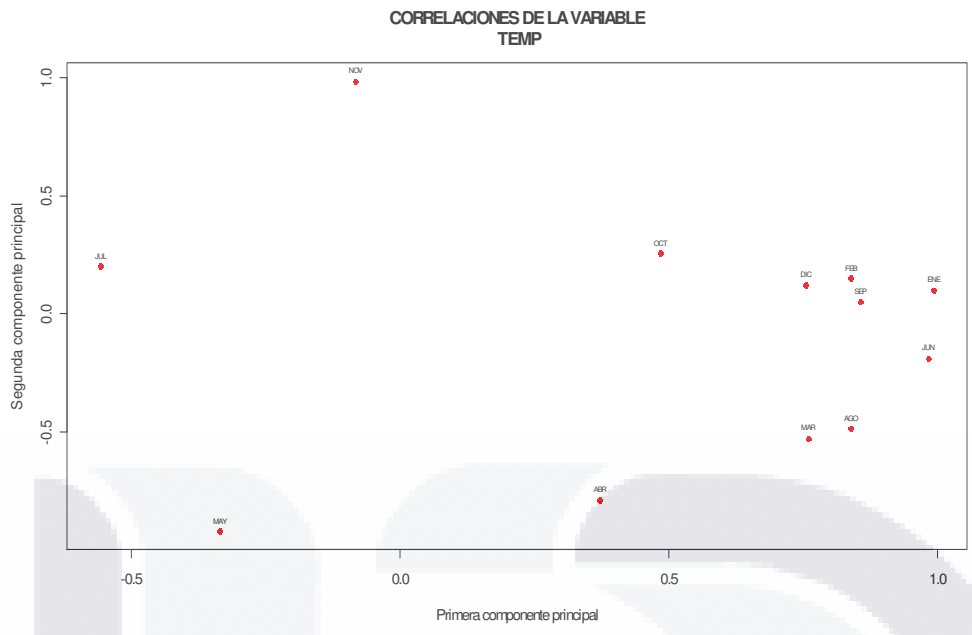


Correlaciones del Óxido de Nitrógeno a través de los meses

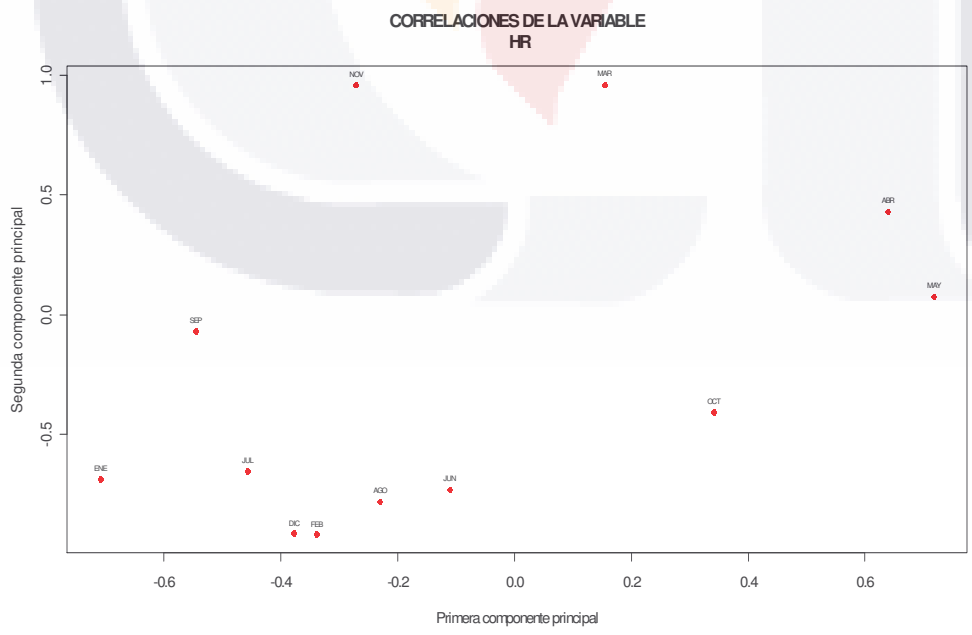
CORRELACIONES DE LA VARIABLE WS



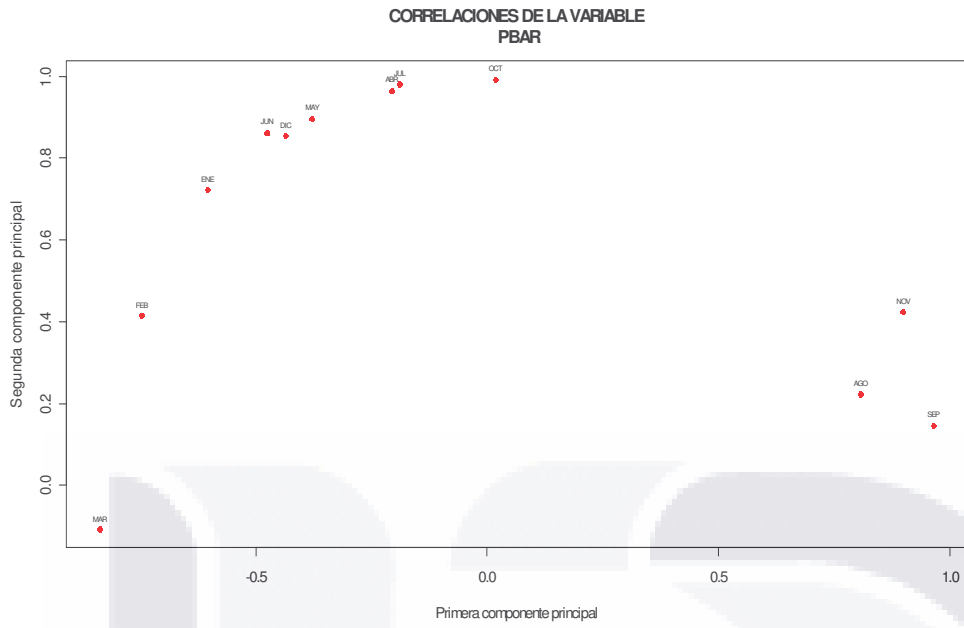
Correlaciones de la Velocidad del Viento a través de los meses



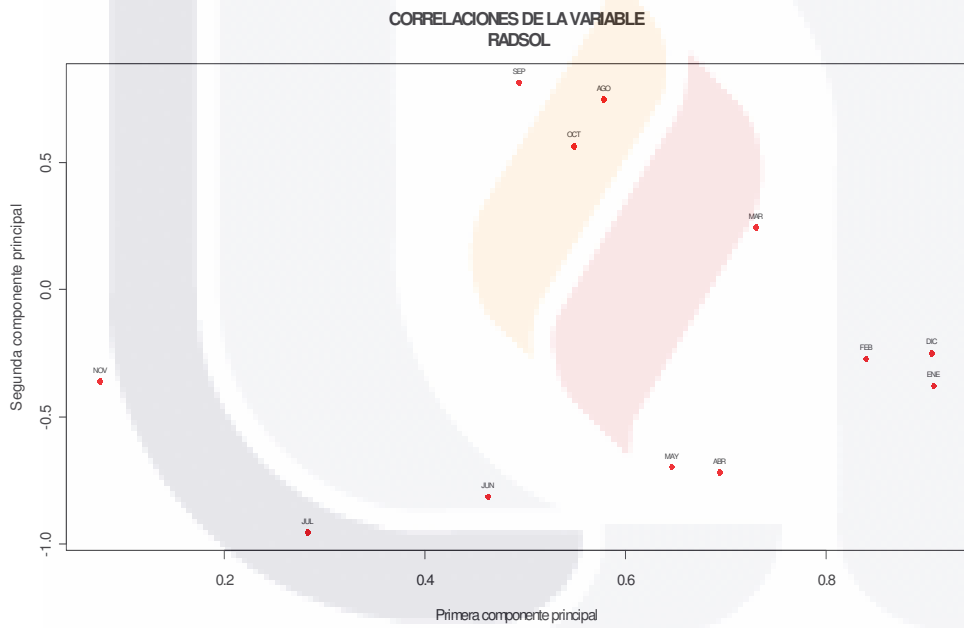
Correlaciones de la Temperatura a través de los meses



Correlaciones de la Humedad Relativa a través de los meses



Correlaciones de la Presión Barométrica a través de los meses



Correlaciones de la Radiación Solar a través de los meses

**TABLAS DE DIFERENCIA EN LOS PUNTAJES DE 2 MESES**

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	MAY	ENE	MAY	PORCENTAJE
<b>Celaya</b>	-0.148	1.525	0.833	2.562	-207.6
<b>Irapuato</b>	-0.527	0.246	0.454	1.282	-182.3
<b>Salamanca</b>	-0.981	-0.734	0.000	0.302	0.0
<b>León</b>	1.655	-1.036	2.636	0.000	100.0

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	JUN	ENE	JUN	PORCENTAJE
Celaya	-0.148	-0.955	0.833	0.000	100.0
Irapuato	-0.527	0.124	0.454	1.079	-137.7
Salamanca	-0.981	-0.752	0.000	0.203	0.0
León	1.655	1.583	2.636	2.539	3.7

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	JUL	ENE	JUL	PORCENTAJE
Celaya	-0.148	-1.030	0.833	0.000	100.0
Irapuato	-0.527	-0.377	0.454	0.653	-43.7
Salamanca	-0.981	-0.248	0.000	0.782	0.0
León	1.655	1.654	2.636	2.684	-1.8

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	AGO	ENE	AGO	PORCENTAJE
Celaya	-0.148	0.756	0.833	1.829	-119.6
Irapuato	-0.527	-1.073	0.454	0.000	100.0
Salamanca	-0.981	1.214	0.000	2.287	0.0
León	1.655	-0.897	2.636	0.176	93.3

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	SEP	ENE	SEP	PORCENTAJE
Celaya	-0.148	-0.686	0.833	0.219	73.7
Irapuato	-0.527	-0.904	0.454	0.000	100.0
Salamanca	-0.981	1.646	0.000	2.550	0.0
León	1.655	-0.056	2.636	0.849	67.8

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	OCT	ENE	OCT	PORCENTAJE
Celaya	-0.148	1.688	0.833	2.605	-212.8
Irapuato	-0.527	-0.917	0.454	0.000	100.0
Salamanca	-0.981	-0.301	0.000	0.616	0.0
León	1.655	-0.469	2.636	0.448	83.0

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	NOV	ENE	NOV	PORCENTAJE
Celaya	-0.148	-1.083	0.833	0.000	100.0
Irapuato	-0.527	-0.503	0.454	0.579	-27.6
Salamanca	-0.981	1.604	0.000	2.687	0.0
León	1.655	-0.019	2.636	1.064	59.6

	PUNTAJES DEL PRIMER COMPONENTE PRINCIPAL		DISTANCIA DEL PUNTAJE RESPECTO AL PUNTAJE MAYOR		REDUCCIÓN DE LAS DISTANCIAS
	ENE	DIC	ENE	DIC	PORCENTAJE
Celaya	-0.148	0.407	0.833	1.720	-106.5
Irapuato	-0.527	-1.313	0.454	0.000	100.0
Salamanca	-0.981	-0.469	0.000	0.844	0.0
León	1.655	1.375	2.636	2.688	-2.0

TABLAS DE MATRICES DE PRODUCTOS ESCALARES POR MES

Abril	Celaya	Irapuato	Salamanca	León
Celaya	8.35	-8.08	-1.42	1.15
Irapuato	-8.08	18.14	-5.11	-4.95
Salamanca	-1.42	-5.11	10.12	-3.59
León	1.15	-4.95	-3.59	7.38

Matriz de Productos Escalares de Abril

Mayo	Celaya	Irapuato	Salamanca	León
Celaya	15.61	0.47	-7.34	-8.74
Irapuato	0.47	4.44	-2.38	-2.52
Salamanca	-7.34	-2.38	11.21	-1.48
León	-8.74	-2.52	-1.48	12.75

Matriz de Productos Escalares de Mayo

Junio	Celaya	Irapuato	Salamanca	León
Celaya	10.46	-0.89	-1.49	-8.08
Irapuato	-0.89	6.60	-4.37	-1.34
Salamanca	-1.49	-4.37	11.69	-5.83
León	-8.08	-1.34	-5.83	15.26

Matriz de Productos Escalares de Junio

Julio	Celaya	Irapuato	Salamanca	León
Celaya	10.51	0.54	-3.19	-7.86
Irapuato	0.54	6.47	-3.20	-3.81
Salamanca	-3.19	-3.20	10.87	-4.47
León	-7.86	-3.81	-4.47	16.15

Matriz de Productos Escalares de Julio

Agosto	Celaya	Irapuato	Salamanca	León
Celaya	9.06	-4.52	0.12	-4.66
Irapuato	-4.52	12.00	-6.23	-1.25
Salamanca	0.12	-6.23	11.57	-5.46
León	-4.66	-1.25	-5.46	11.37

Matriz de Productos Escalares de Agosto

Septiembre	Celaya	Irapuato	Salamanca	León
Celaya	10.41	-2.30	-7.07	-1.04
Irapuato	-2.30	11.85	-7.83	-1.72
Salamanca	-7.07	-7.83	16.94	-2.04
León	-1.04	-1.72	-2.04	4.80

Matriz de Productos Escalares de Septiembre

Octubre	Celaya	Irapuato	Salamanca	León
Celaya	14.80	-6.00	-4.43	-4.36
Irapuato	-6.00	10.74	-3.82	-0.92
Salamanca	-4.43	-3.82	10.71	-2.46
León	-4.36	-0.92	-2.46	7.74

Matriz de Productos Escalares de Octubre

Noviembre	Celaya	Irapuato	Salamanca	León
Celaya	11.63	0.41	-7.26	-4.78
Irapuato	0.41	6.08	-4.89	-1.61
Salamanca	-7.26	-4.89	16.02	-3.87
León	-4.78	-1.61	-3.87	10.27

Matriz de Productos Escalares de Noviembre



Diciembre	Celaya	Irapuato	Salamanca	León
Celaya	6.93	-3.04	-4.84	0.95
Irapuato	-3.04	13.78	-0.29	-10.44
Salamanca	-4.84	-0.29	9.47	-4.33
León	0.95	-10.44	-4.33	13.82

Matriz de Productos Escalares de Diciembre

### TABLAS DE LA INTERESTRUCTURA Wk

Abril	Celaya	Irapuato	Salamanca	León
Celaya	2.09	-2.02	-0.36	0.29
Irapuato	-2.02	4.54	-1.28	-1.24
Salamanca	-0.36	-1.28	2.53	-0.90
León	0.29	-1.24	-0.90	1.85

Interestructura de Abril de 2006

Mayo	Celaya	Irapuato	Salamanca	León
Celaya	3.90	0.12	-1.84	-2.19
Irapuato	0.12	1.11	-0.60	-0.63
Salamanca	-1.84	-0.60	2.80	-0.37
León	-2.19	-0.63	-0.37	3.19

Interestructura de Mayo de 2006

Junio	Celaya	Irapuato	Salamanca	León
Celaya	2.61	-0.22	-0.37	-2.02
Irapuato	-0.22	1.65	-1.09	-0.34
Salamanca	-0.37	-1.09	2.92	-1.46
León	-2.02	-0.34	-1.46	3.81

Interestructura de Junio de 2006

Julio	Celaya	Irapuato	Salamanca	León
Celaya	2.63	0.14	-0.80	-1.97
Irapuato	0.14	1.62	-0.80	-0.95
Salamanca	-0.80	-0.80	2.72	-1.12
León	-1.97	-0.95	-1.12	4.04

Interestructura de Julio de 2006

Agosto	Celaya	Irapuato	Salamanca	León
Celaya	2.26	-1.13	0.03	-1.16
Irapuato	-1.13	3.00	-1.56	-0.31
Salamanca	0.03	-1.56	2.89	-1.36
León	-1.16	-0.31	-1.36	2.84

Interestructura de Agosto de 2006

Septiembre	Celaya	Irapuato	Salamanca	León
Celaya	2.60	-0.58	-1.77	-0.26
Irapuato	-0.58	2.96	-1.96	-0.43
Salamanca	-1.77	-1.96	4.23	-0.51
León	-0.26	-0.43	-0.51	1.20

Interestructura de Septiembre de 2006

Octubre	Celaya	Irapuato	Salamanca	León
Celaya	3.70	-1.50	-1.11	-1.09
Irapuato	-1.50	2.69	-0.96	-0.23
Salamanca	-1.11	-0.96	2.68	-0.62
León	-1.09	-0.23	-0.62	1.94

Interestructura de Octubre de 2006

Noviembre	Celaya	Irapuato	Salamanca	León
Celaya	2.91	0.10	-1.81	-1.20
Irapuato	0.10	1.52	-1.22	-0.40
Salamanca	-1.81	-1.22	4.01	-0.97
León	-1.20	-0.40	-0.97	2.57

Interestructura de Noviembre de 2006

Diciembre	Celaya	Irapuato	Salamanca	León
Celaya	1.73	-0.76	-1.21	0.24
Irapuato	-0.76	3.44	-0.07	-2.61
Salamanca	-1.21	-0.07	2.37	-1.08
León	0.24	-2.61	-1.08	3.46

Interestructura de Diciembre de 2006

**GRÁFICAS DE LA MATRIZ COMPROMISO POR MES**

Abril	Celaya	Irapuato	Salamanca	León
Celaya	0.69	-0.67	-0.12	0.10
Irapuato	-0.67	1.50	-0.42	-0.41
Salamanca	-0.12	-0.42	0.83	-0.30
León	0.10	-0.41	-0.30	0.61

Matriz Compromiso de Abril

Mayo	Celaya	Irapuato	Salamanca	León
Celaya	1.45	0.04	-0.68	-0.81
Irapuato	0.04	0.41	-0.22	-0.23
Salamanca	-0.68	-0.22	1.04	-0.14
León	-0.81	-0.23	-0.14	1.19

Matriz Compromiso de Mayo

Junio	Celaya	Irapuato	Salamanca	León
Celaya	0.97	-0.08	-0.14	-0.75
Irapuato	-0.08	0.61	-0.41	-0.12
Salamanca	-0.14	-0.41	1.08	-0.54
León	-0.75	-0.12	-0.54	1.42

Matriz Compromiso de Junio

Julio	Celaya	Irapuato	Salamanca	León
Celaya	0.98	0.05	-0.30	-0.74
Irapuato	0.05	0.61	-0.30	-0.36
Salamanca	-0.30	-0.30	1.02	-0.42
León	-0.74	-0.36	-0.42	1.51

Matriz Compromiso de Julio

Agosto	Celaya	Irapuato	Salamanca	León
Celaya	0.79	-0.40	0.01	-0.41
Irapuato	-0.40	1.05	-0.55	-0.11
Salamanca	0.01	-0.55	1.01	-0.48
León	-0.41	-0.11	-0.48	1.00

Matriz Compromiso de Agosto

Septiembre	Celaya	Irapuato	Salamanca	León
Celaya	0.91	-0.20	-0.62	-0.09
Irapuato	-0.20	1.04	-0.69	-0.15
Salamanca	-0.62	-0.69	1.48	-0.18
León	-0.09	-0.15	-0.18	0.42

Matriz Compromiso de Septiembre

Octubre	Celaya	Irapuato	Salamanca	León
Celaya	1.28	-0.52	-0.38	-0.38
Irapuato	-0.52	0.93	-0.33	-0.08
Salamanca	-0.38	-0.33	0.92	-0.21
León	-0.38	-0.08	-0.21	0.67

Matriz Compromiso de Octubre

Noviembre	Celaya	Irapuato	Salamanca	León
Celaya	1.09	0.04	-0.68	-0.45
Irapuato	0.04	0.57	-0.46	-0.15
Salamanca	-0.68	-0.46	1.50	-0.36
León	-0.45	-0.15	-0.36	0.96

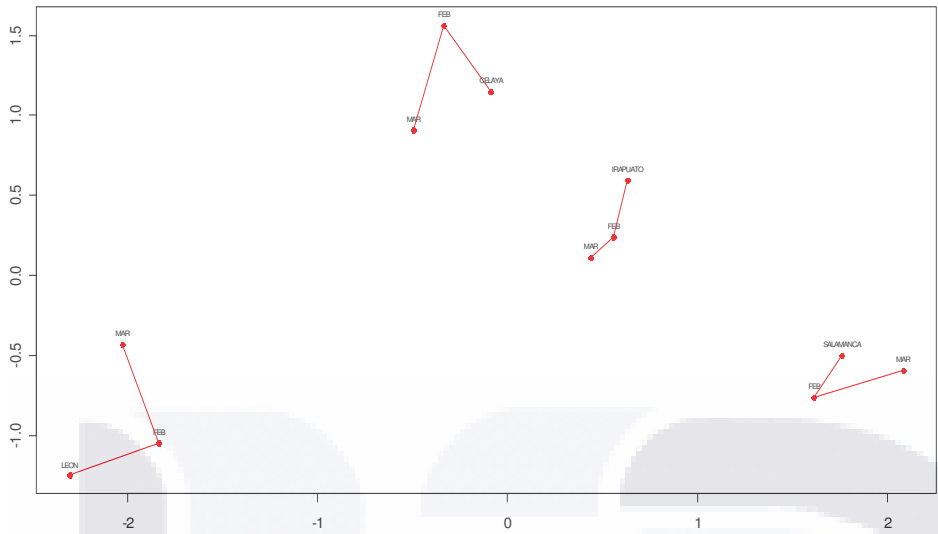
Matriz Compromiso de Noviembre

Diciembre	Celaya	Irapuato	Salamanca	León
Celaya	0.62	-0.27	-0.43	0.09
Irapuato	-0.27	1.23	-0.03	-0.94
Salamanca	-0.43	-0.03	0.85	-0.39
León	0.09	-0.94	-0.39	1.24

Matriz Compromiso de Diciembre

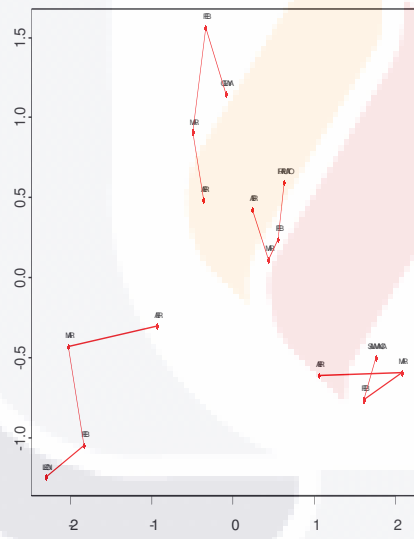
GRÁFICAS DE LAS TRAYECTORIAS POR MES

TRAYECTORIAS DE LAS CIUDADES DEL CORREDOR INDUSTRIAL DE ENE A MAR DE 2006



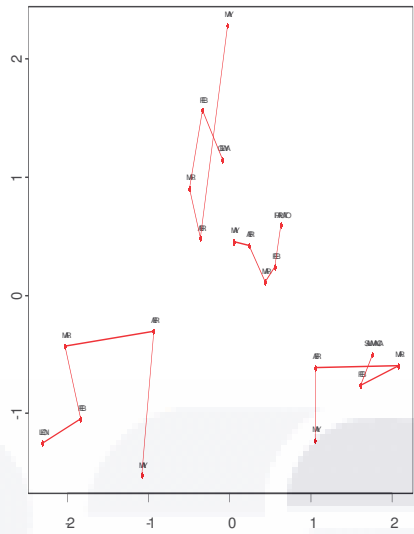
Trayectorias hasta Marzo

TRAYECTORIAS DE LAS CIUDADES DEL CORREDOR INDUSTRIAL DE ENE A MARZO DE 2006



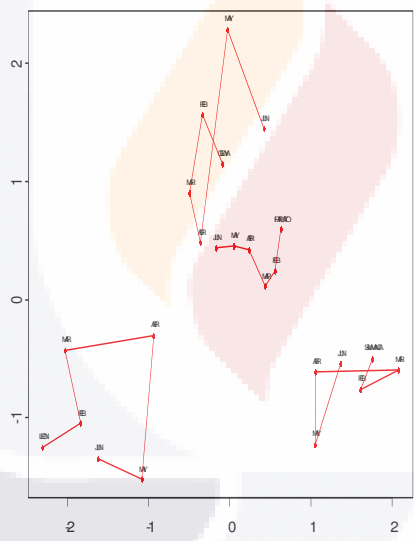
Trayectorias hasta Abril

**TRAYECTORIAS DE LOS INDICADORES DE RIESGO**  
**DE MAYO**



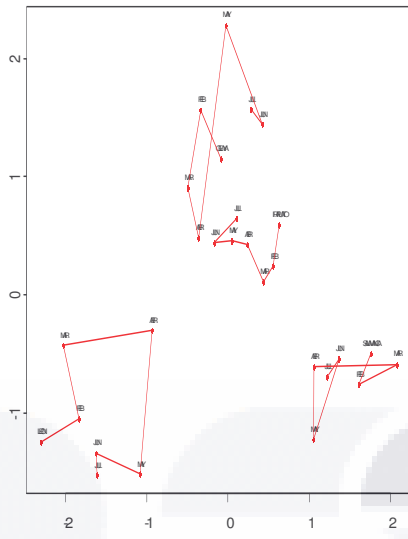
Trayectorias hasta Mayo

**TRAYECTORIAS DE LOS INDICADORES DE RIESGO**  
**DE JUNIO**



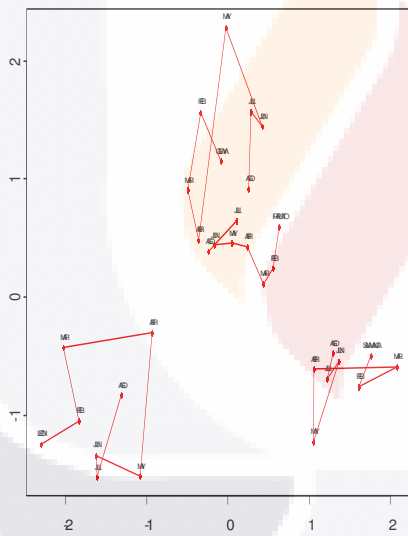
Trayectorias hasta Junio

**TRAYECTORIAS DE LOS COORDENADOS  
DEPARTAMENTO**



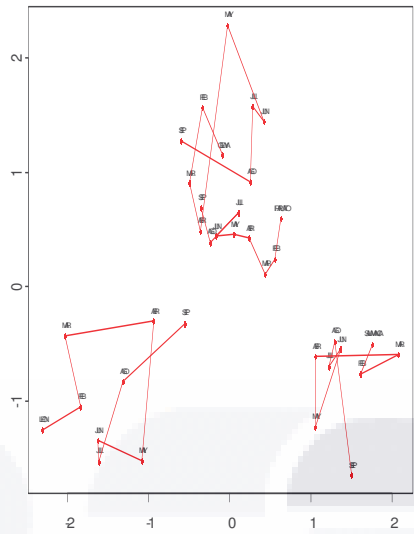
Trayectorias hasta Julio

**TRAYECTORIAS DE LOS COORDENADOS  
DEPARTAMENTO**



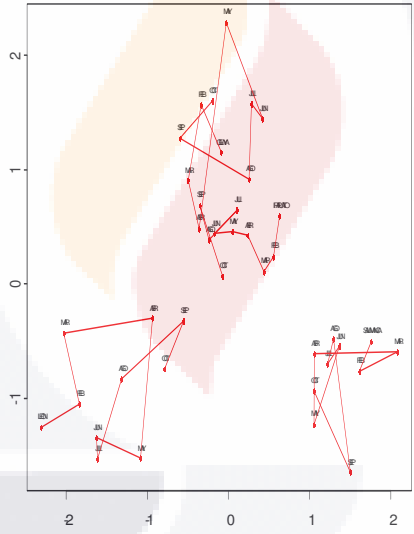
Trayectorias hasta Agosto

**FACTORES ANÁLISIS DE CLUSTRO DE PA**  
**DIAS 4/2/2015**



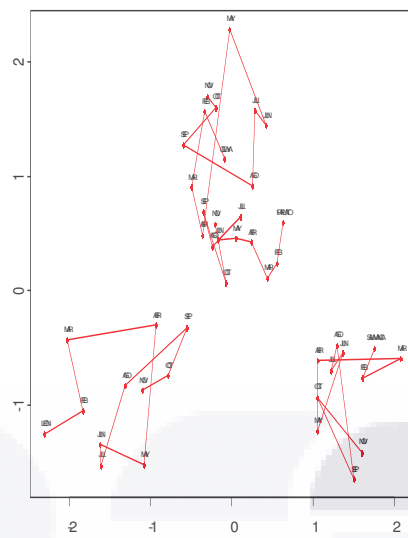
Trayectorias hasta Septiembre

**FACTORES ANÁLISIS DE CLUSTRO DE PA**  
**DIAS 4/2/2015**



Trayectorias hasta Octubre

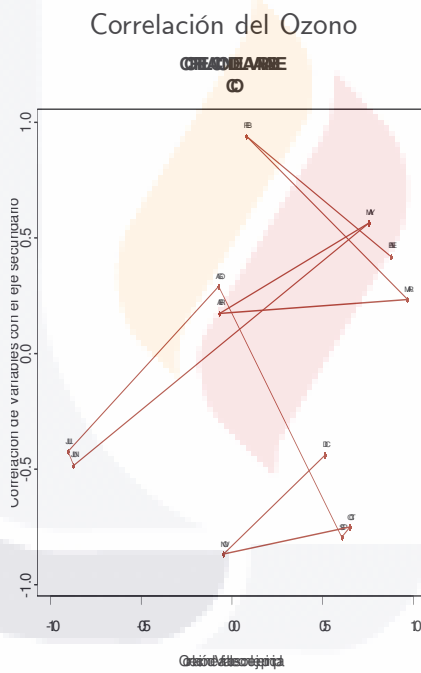
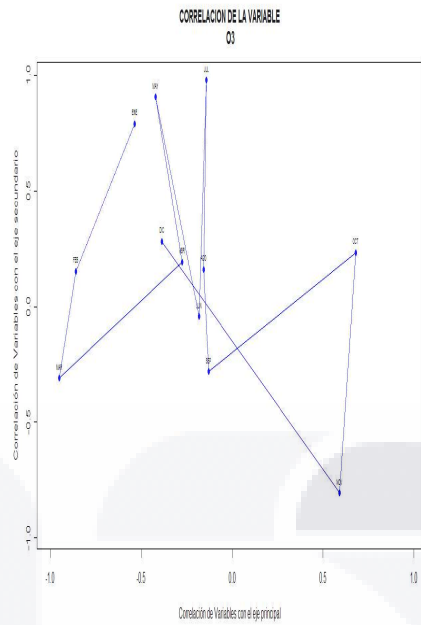
**TRAYECTORIAS ANIMALES EN CHIRIQUÍ**  
**DIRECCIONES**



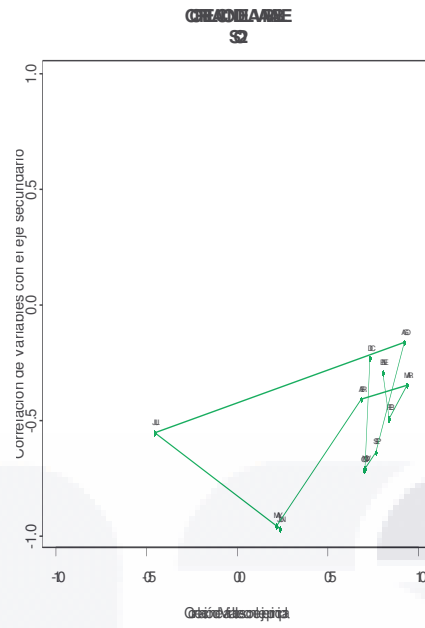
Trayectorias hasta Noviembre



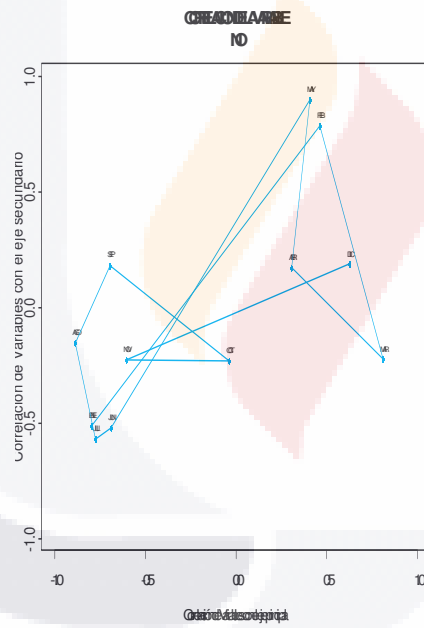
GRÁFICAS DE LAS CORRELACIONES POR VARIABLES



Correlación del Monóxido de Carbono

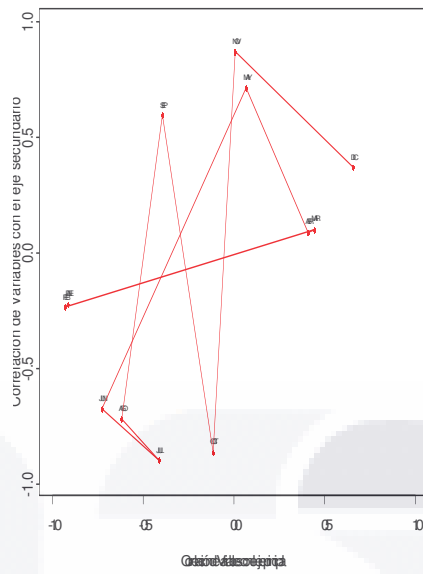


Correlación del Dióxido de Azufre



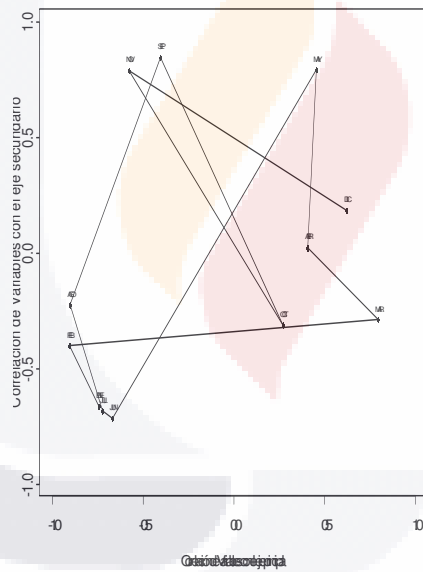
Correlación del Monóxido de Nitrógeno

**GRACIA LA ARBE**  
**NO2**

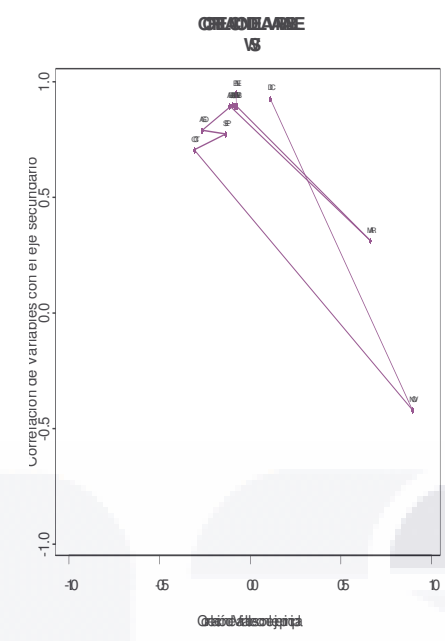


Correlación del Dióxido de Nitrógeno

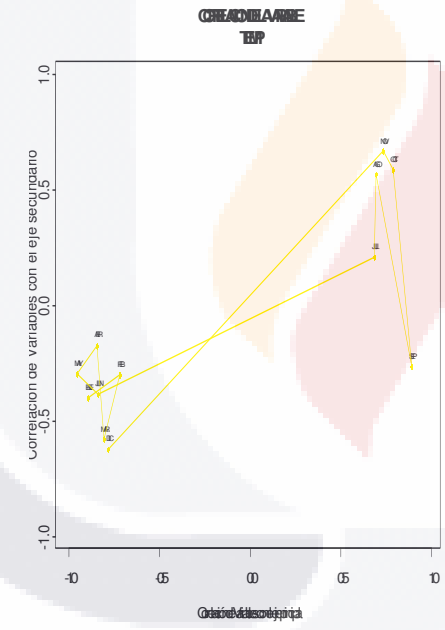
**GRACIA LA ARBE**  
**NOx**



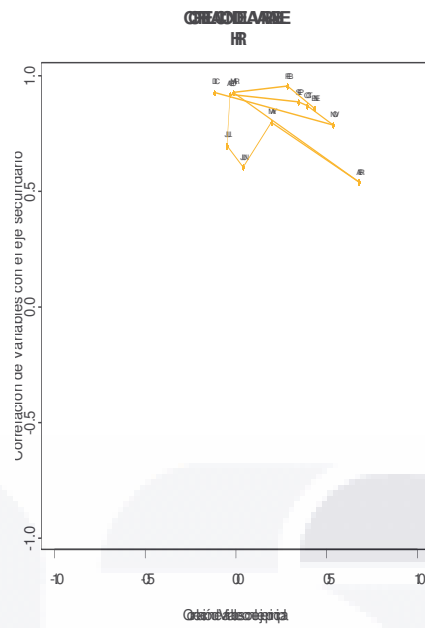
Correlación del Óxido de Nitrógeno



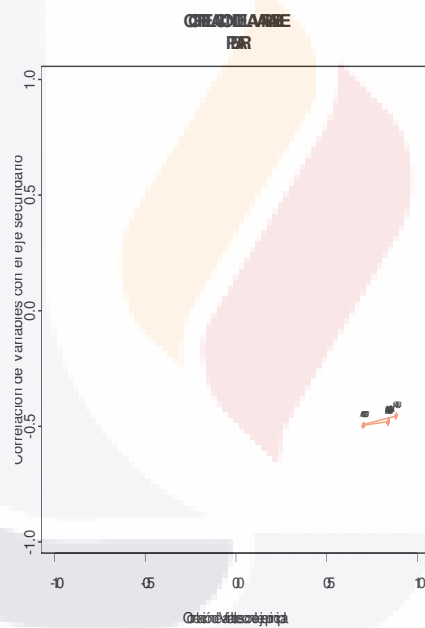
Correlación de la Velocidad del Viento



Correlación de la Temperatura

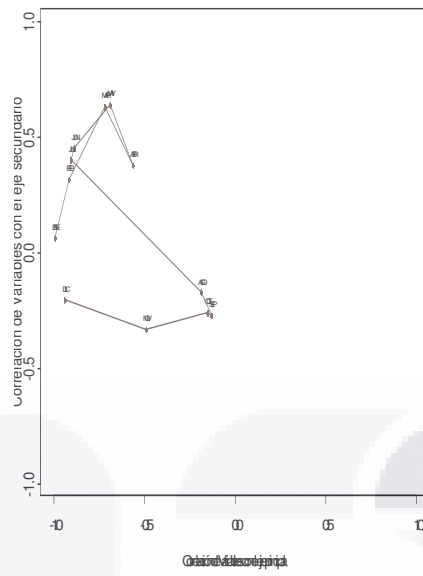


Correlación de la Humedad Relativa



Correlación de la Presión Barométrica

GRÁFICO 11  
FASD



Correlación de la Radiación Solar

# Bibliografía y Referencias

- [AHVD2007] Abdi Hervé, Valentin Dominique (2007) *The STATIS Method*.
- [BBC2009] Bárcenas Blancarte, C. (2009) *Red de Monitoreo de Calidad del Aire del Estado de Guanajuato*, Presentación en el Simposium Internacional de Gestión de Calidad del Aire, León, Gto. URL [http://ecologia.guanajuato.gob.mx/2009/archivos/simposium\\_abril/10.pdf](http://ecologia.guanajuato.gob.mx/2009/archivos/simposium_abril/10.pdf)
- [CAG2009] *Red de Monitoreo de Calidad del Aire del Estado de Guanajuato* (2009) URL <http://201.120.52.214/calidadaire/usuario/index.php>
- [CEGV1998] Castillo Elizondo W., González Varela J. (1998) *STATIS Dual: Software y Análisis de Datos Reales*.
- [DCA2009] *Definición de Contaminación Atmosférica* (2009) URL <http://es.wikipedia.org/wiki/Contaminación>
- [EPL1986] E. Pla, Laura (1986) *Análisis Multivariado: Método de Componentes Principales*.
- [GRF2008] Geyne Rajme, Frida (2008) *Análisis Longitudinal de la Marginación en las Entidades Federativas de México (1990-2005)*.
- [LCE1994] Lavit Christine, Escoufier Yves, et. al. (1994) *The ACT (STATIS method)*.
- [LAR2007] López A. J. G., Arredondo R. J., Robledo B. D. (2007) *La Calidad del Aire en Guanajuato. Informe 2006 de estado y de tendencias*.
- [LGE2009] *Ley General del Equilibrio Ecológico y la Protección al Ambiente* (2009) URL <http://www.diputados.gob.mx/LeyesBiblio/pdf/148.pdf>
- [PACR2006] Pérez-Abreu Carrión R. A. (2006) *Estudio de la Calidad del Aire en el Estado de Guanajuato Mediante el Diseño de un Modelo Estadístico de Diagnóstico y Pronóstico*
- [PEAJ2003] Paradis Emmanuel, Ahumada Jorge A. (2003) *R para principiantes*, URL [http://cran.r-project.org/doc/contrib/rdebuts\\_es.pdf](http://cran.r-project.org/doc/contrib/rdebuts_es.pdf)
- [RPSC2009] *The R Project for Statistical Computing* (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.