

Seminário Latinoamericano:
Instrumentos y metodologias para un
observatório de Clima y su impacto en la
salud humana

*Sergio Ibarra-Espinosa, Universidade de São Paulo,
sergio.ibarra@usp.br, ibarraespinosa.github.io*

2019-09-11

Contents

Curso de R, contaminacion atmosferica y mas	5
Aprender Git	5
Clonar este contenido	5
1 Sistemas de informacion con datos de salud en Chile	7
1.1 Encuesta Nacional de Salud (ENS)	7
1.2 Departamento de Estadisticas e Informaciones de Salud	8
1.3 Epidemiologia	9
1.4 Encuesta de caracterizacion socioeconomica (CASEN)	9
1.5 Estadisticas generales	9
1.6 CEPAL STAT	9
1.7 Banco Interamericano de Desarrollo	9
1.8 Datos de contaminacion atmosferica y meteorologia	9
1.9 Algunos ejemplos de datos de series de tiempo	10
2 Impacto de las emisiones antropogénicas en la calidad del aire y clima	13
2.1 Contaminacion atmosferica - Introduccion	14
2.2 Contaminantes atmosfericos	16
2.3 Forzantes climaticos y gases de efecto invernadero	22
2.4 Emisiones y sus fuentes	32
2.5 Nuevos modelos de emisiones: VEIN	36
2.6 Nuevos datos, satelites y modelacion inversa	36
3 Taller VECTORES: Aplicación de software de información geográfica y modelado	43
3.1 Instalacion	44
3.2 R desde Excel, libreoffice, archivos de texto, etc.	44
3.3 Vectores	48
3.4 Geoprocесamiento	53
4 Taller RASTER Y CUBOS DE DATOS VECTORIALES: Aplicación de software de información geográfica y modelado	59

5 Estudios de series de tiempo para asociar salud y variables ambientales	73
5.1 Distribucion	73
5.2 Efecto de la contaminacion	74
5.3 Efecto de la temperatura	75

Curso de R, contaminacion atmosferica y mas

Este curso online contendra las siguientes informaciones

- Sistemas de informacion con datos de salud en Chile (gracias Paty Matus)
- Impacto de las emisiones antropogenicas en la salud y clima
- R desde Excel
- Leer y procesar vectores espaciales con **sf** (Pebesma, 2018)
- Leer y procesar informacion en grillas espaciales (raster) con stars(Pebesma, 2019) y raster(Hijmans, 2019)

Aprender Git

Para aprender GIT puedes ver:

- <https://git-scm.com/book/es/v1/Empezando>
- <https://learngitbranching.js.org/>
- <https://try.github.io/>

Clonar este contenido

Para clonar este contenido haz:

```
git clone https://github.com/ibarraespinosa/UBA.git
```


Chapter 1

Sistemas de informacion con datos de salud en Chile

- Sistema de información en salud existentes
- Enfasis en las fuentes de información y las escala temporal/espacial que manejan
- Series de tiempo disponible por fuente
- Instituciones a cargo de la captura, procesamiento y análisis
- Disponibilidad de los datos e indicadores que producen
- Otros

1.1 Encuesta Nacional de Salud (ENS)

La ENS es una encuesta realizada por el Ministerio de Salud para identificar cuales son las enfermedades que sufren y los tratamientos que reciben todas las personas con mas de 15 años que viven en Chile. De esta forma es posible es posible realizar diagnosticos, identificar problemas y formular politicas planes y proyectos para mejor la salud de las personas.

- *Organismo responsable:* Ministerio de Salud, Departamento de Epidemiología
Gobierno de Chile.
- *Organismo ejecutor:* Pontificia Universidad Católica de Chile (PUC).
- *Población objetivo:* Personas de 15 años y más, chilenas o extranjeras que residen habitualmente en viviendas particulares ocupadas, localizadas en zonas urbanas y rurales de las quince regiones de Chile.

8CHAPTER 1. SISTEMAS DE INFORMACION CON DATOS DE SALUD EN CHILE

- *Representatividad:* Nacional, regional y Urbano/Rural.
- *Modo de aplicación:* Entrevista personal en hogar (Sistema de captura electrónica: Tablet), aplicada por encuestador y profesional enfermera de acuerdo al tipo de cuestionario.
- *Período de trabajo de campo:* Agosto 2016 a marzo 2017
- *Tamaño muestral:* 6.233 encuestados, de los cuales 5.520 cuentan con exámenes de laboratorio de acuerdo a protocolo. 37,1% hombres, 62,9% mujeres.
- *Error muestral:* Error absoluto de muestreo de 2,6% a nivel nacional, raíz del efecto de diseño de 1,797, estimaciones con 95% de confianza y error relativo inferior a 30%.

Algunos resultados:

- Consumo de tabaco: 66,7% no fuma, 33,% fuma.
- Consumo riesgoso de alcohol 11,7%, 20,5% hombres, 3,3% mujeres.
- Sedentarismo: 86,7%, 83,3% hombre, 90,0% mujeres.
- Estado nutricional: 1,3% enfraquecido, 24,5% normal, 39,8% sobrepeso, 31,2% obeso, 3,2% obeso morbido.
- Sospecha de hipertension: 27,6%.
- Sospecha de diabetes: 12,3%.
- Autoreporte de infarto agudo al miocardio: 3,3%.
- Autoreporte de ataque cerebro vascular: 2,6%.

Fuentes:

- https://www.minsal.cl/wp-content/uploads/2017/11/ENS-2016-17_PRIMEROS-RESULTADOS.pdf
- https://www.minsal.cl/wp-content/uploads/2018/01/2-Resultados-ENS_MINSAL_31_01_2018.pdf
- <http://www.encuestas.uc.cl/ens/index.html>

1.2 Departamento de Estadísticas e Informaciones de Salud

- Resúmenes estadísticos mensuales (REM). Vea el manual
- Defunciones
- Egresos
- Nacimientos
- Atenciones de urgencia
- Enfermedades de notificación obligatoria
- Enfermedades transmitidas por alimentos

- Tuberculosis

1.3 Epidemiologia

- Vigilancia Epidemiologica

1.4 Encuesta de caracterizacion socioeconomica (CASEN)

“La Encuesta de Caracterización Socioeconómica Nacional (Casen) del Ministerio de Desarrollo Social es una encuesta a hogares, de carácter multipropósito, es decir, que abarca diversos temas como educación, trabajo, ingresos, salud, entre otros; además es una encuesta transversal, por lo tanto, incluye a todo el espectro de la población del país.”

1.5 Estadisticas generales

- Instituto Nacional de Estadisticas

1.6 CEPAL STAT

- Estadisticos e indicadores
- Perfiles Nacionales
- Publicaciones y estadisticas

1.7 Banco Interamericano de Desarrollo

- Educacion
- Mercado Laboral
- Ingreso
- Pobreza
- Demografia

1.8 Datos de contaminacion atmosferica y meteorologia

- SINCA
- QUALAR
- RESEARCH DATA ARCHIVE
- CMIP
- AchemMip

- IPCC DATA
- RCP
- TELECONNECTIONS

1.9 Algunos ejemplos de datos de series de tiempo

Efecto de la contaminación atmosférica sobre los accidentes cerebrovasculares

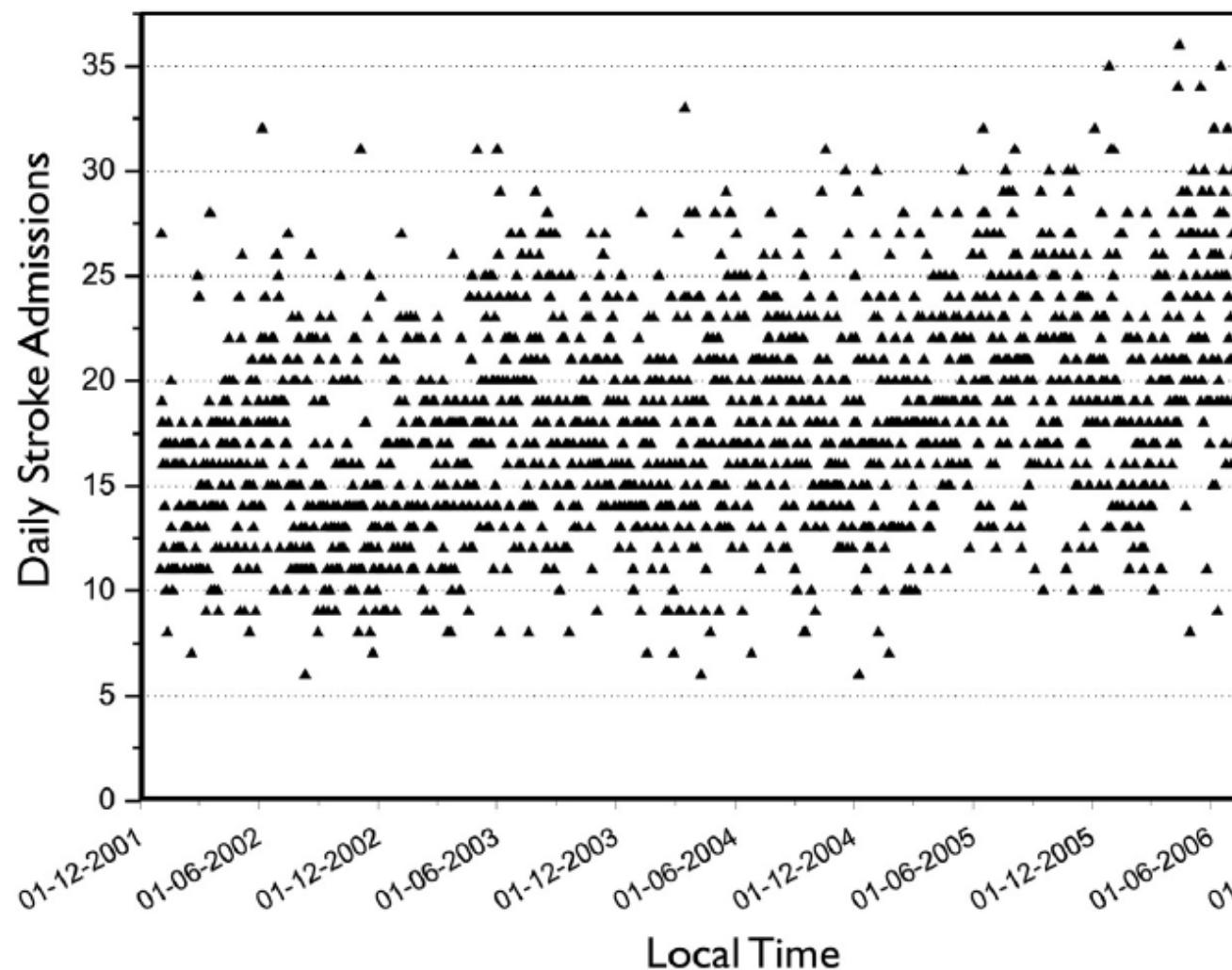


Fig. 2. The daily stroke admissions in Santiago during the study period.

Figure 1.1: Accidentes Cerebro Vasculares Chile (Leiva et al., 2013)

12CHAPTER 1. SISTEMAS DE INFORMACION CON DATOS DE SALUD EN CHILE

Chapter 2

Impacto de las emisiones antropogénicas en la calidad del aire y clima

La contaminacion es una amenaza a la salud global que causo 9 millones de muertes en 2015, 16% de las muertes (Landrigan et al., 2018). Ademas, existe consenso cientifico de la existencia y relacion con la actividad humana sobre el cambio climatico (Cook et al., 2016). Tanto la contaminacion atmosferica como los gases de efecto invernadero, conocidos tambien como forzantes climaticos de vida corta y larga son causados por compuestos quimicos liberados en la atmosfera, las llamadas emisiones. Por lo tanto, es importante realziar y entender la caracterizacion espacial y temporal de las emisiones, asi como sus efectos en la salud y en el clima.

Algunos efectos de la contaminacion atmosferica en la salud

- incrementa la morbilidad, natalidad
- aumenta la tasa de nacimientos con problemas (menos peso, inteligencia)
- incrementa riesgo de cancer
- etc et etc

La atmosfera es una delgada capa sobre la tierra, 50% de su masa esta 5.6 km esta compuestos por varios gases (Brasseur and Jacob, 2017), cuya composicion es:

Gas	Razon molar $mol \cdot mol^{-1}$	Principal fuente y comentarios (Brasseur et al., 1999)
Nitrogeno (N_2)	0.78	Biologica

Gas	Razon molar $mol \cdot mol^{-1}$	Principal fuente y comentarios (Brasseur et al., 1999)
Oxigeno (O_2)	0.21	Biologica
Argon (A_r)	0.0093	Inerte
Dioxido de carbono (CO_2)	$400 \cdot 10^{-6}$	Combustion, oceano, biosfera
Neon (N_e)	$18 \cdot 10^{-6}$	Inerte
Ozono (O_3)	$0.01 - 10 \cdot 10^{-6}$	Fotoquimico
Helio (H_e)	$5.2 \cdot 10^{-6}$	Inerte
Metano (CH_4)	$1.8 \cdot 10^{-6}$	Biogenico y antropogenico
Hydrogeno (H_2)	$500 \cdot 10^{-9}$	Biogenico, antropogenico y fotoquimico
Oxido nitroso (N_2O)	$330 \cdot 10^{-9}$	Biogenico y antropogenico

2.1 Contaminacion atmosferica - Introduccion

“La contaminación del aire es un determinante importante de la salud. La OMS estima que en 2012 alrededor de 1 de cada 8 muertes se atribuyeron a la exposición a la contaminación del aire, lo que lo convierte en el mayor factor de riesgo ambiental para la mala salud.”(OMS, 2019)

La ciencia de la contaminacion atmosferica, si bien reciente, ha sido desarrollada debido a los avances de en la comprension de la meteorologia. Problemas relacionados con la contaminacion atmosferica han sido descritos en obras literarias y cartas a lo largo de la historia. Por ejemplo, se cree que el primer caso reportado sobre los efectos de la contaminacion atmosférica en la salud es sobre Gaius Plinius Secundus, Geografo, (AD 23-AD 79), quien habria fallecido los efectos de la **emisiones** del volcan Vesuvius (Pulses and Helsinga, 2010; Wikipedia contributors, 2019d). La erupcion del volcan Vesuvius duro 19 horas, con altura de lacolumna entre 14 y 32 km y deposicion de material piroplastico de hasta $2500 kg \cdot m^{-2}$ (Macedonio et al., 1988).

Sin embargo han sido los grandes episodios de contaminacion los que han gatillado su estudio y gestion por parte de los tomadores de decisiones. Entre ellos se pueden mencionar el desastre de Londres 1952 y la acidificacion de los lagos escandinavos.



Figure 2.1: An eruption of Vesuvius seen from Portici, by Joseph Wright (ca. 1774-6), Dominio Publico

2.1.1 El desastre de Londres 1952

Altas concentraciones de contaminantes fueron ocurrieron entre el 5 y 9 de Diciembre de 1952 en Londres, Inglaterra. Para comparacion, vea el monumento “Columna de Nelson” en condiciones normales(Wikipedia contributors, 2019e), y el dia de la llamada “Gran Niebla de Londres”(Wikipedia contributors, 2019b)

2.1.1.1 Consecuencias

A pesar que durante la fecha, las autoridades no consideraron el efecto de la contaminacion, este evento si causoo gran impacto en la comunidad (Wikipedia contributors, 2019c). Estudios posteriores cuantificaron un impacto en **12.000 muertos** asociados a este episodio de contaminacion (Bell and Davis, 2001).

comentarios?

2.1.1.2 CLRTAP

Tarea / Tema de casa / Homework:

[1] "Investigue que es, causas, consecuencias y poltiica ambiental asociada a CLRTA"

2.2 Contaminantes atmosfericos

Las emisiones liberadas a la atmosfera impactan la salud, meteorologia y clima en diferentes escalas como se ve en la siguiente figura.

Por lo tanto para entender las emisiones necesitamos responder las siguientes preguntas:

- Que?
- Como?
- Cuando?
- Donde?

Los contaminantes atmosfericos suelen ser clasificados como:

- Primarios: emitidos directamente en la atmosfera. Ejemplo: CO.
- Secundarios: formados en la atmosfera. Ejemplo: O_3 .

Existen muchos contaminantes atmosfericos de interes cientifico como

- Ozono O_3
- Monoxido de Carbono CO
- Radicales de Oxidos de nitrogeno $NO_X \equiv NO + NO_2$
- Compuestos organicos volatiles COV
- Radicales de Halogeno
- Especies de azufre SO_2 , SO_4 , H_2SO_4
- Aersoles

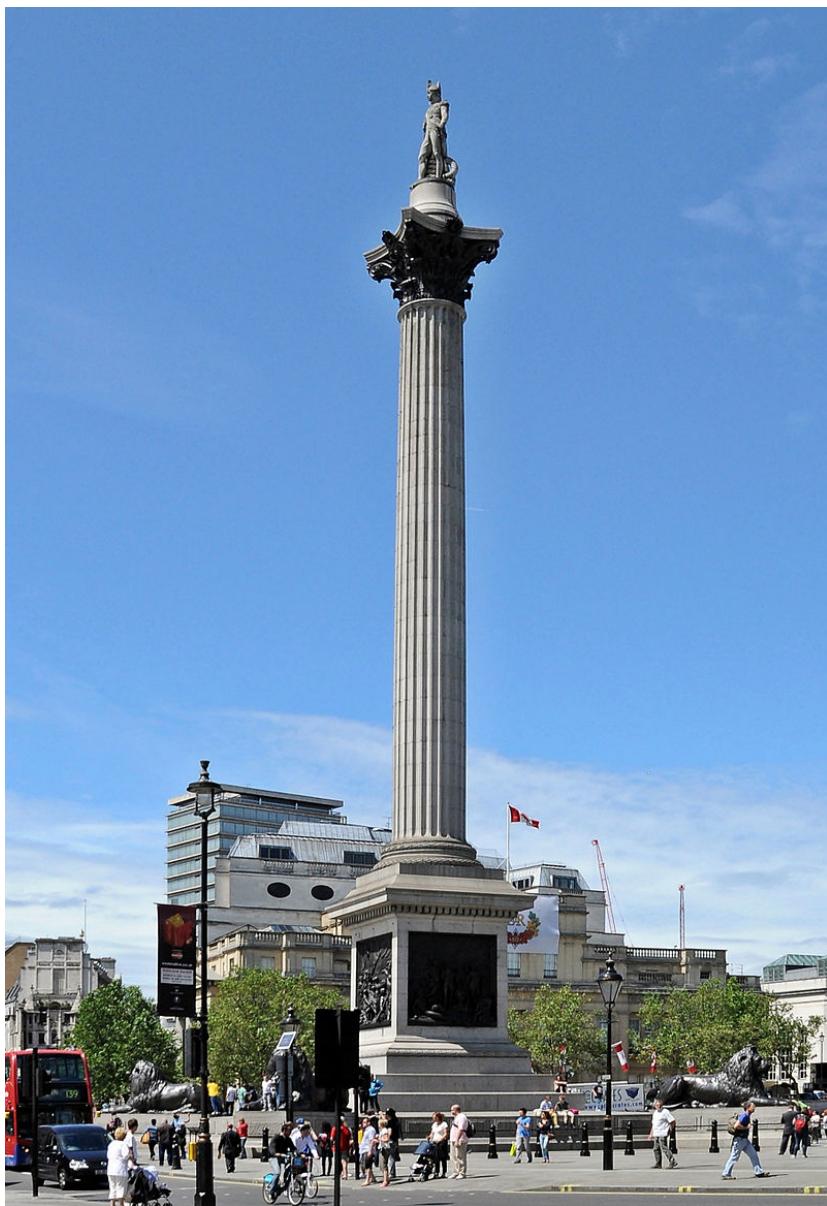


Figure 2.2: Columna de Nelson, Dominio Publico



Figure 2.3: Columna de Nelson durante la Gran Niebla de 1952, Dominio Público

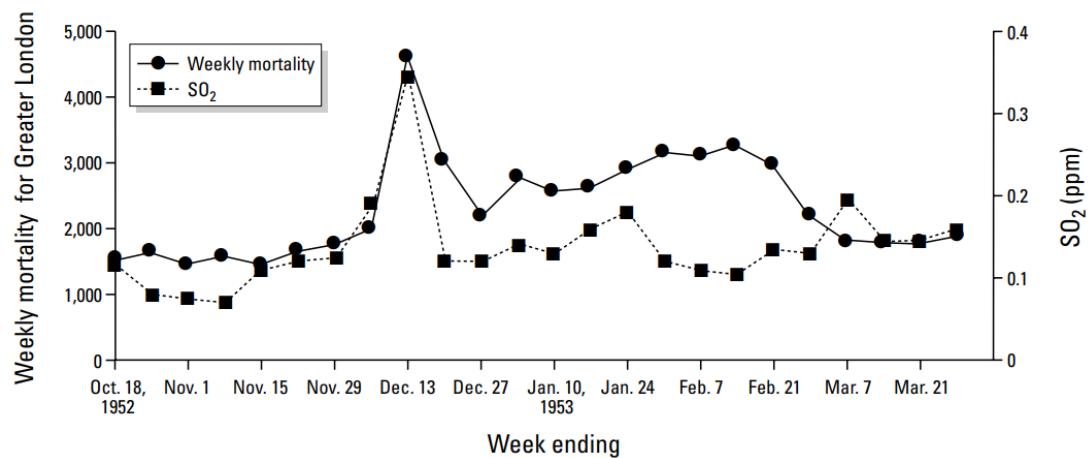


Figure 1. Approximate weekly mortality and SO₂ concentrations for Greater London, 1952–1953.

Figure 2.4: Mortalidad semanal y concentraciones de SO₂ en Londres 1952 (Bell and Davis, 2001)

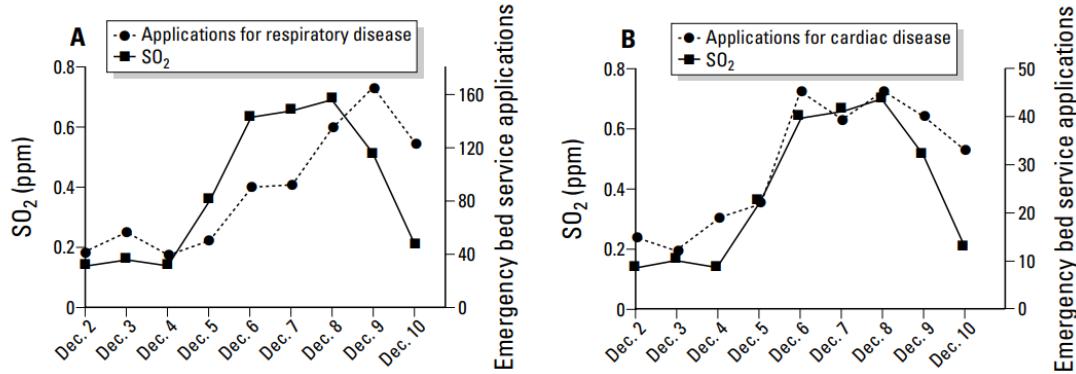


Figure 4. (A) SO_2 and emergency bed service applications for respiratory disease for Greater London, 1952. (B) SO_2 and emergency bed service applications for cardiac disease for Greater London, 1952.

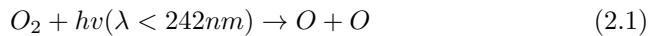
Figure 2.5: Mortalidad semanal y concentraciones de SO_2 en Lobres 1952 (Bell and Davis, 2001)

Cada contaminante tiene un extenso cuerpo teórico. En esta parte del curso nos referiremos brevemente al O_3 , aerosoles y especies de azufre, sin embargo, el estudiante puede revisar la bibliografía para se profundizar.

2.2.1 Ozono O_3

Respecto del ozono es necesario mencionar que existen dos tipos, el estratosférico (bueno) y el troposférico (malo), como es explicado por Brasseur and Jacob (2017).

El ciclo de ozono troposférico comienza con la fotolisis de una molécula de oxígeno debido a radiación solar de longitud de onda mínima 242 nm. Posteriormente, un átomo de oxígeno se combina con otra molécula de oxígeno para formar ozono. La molécula M (cuerpo inerte, como N_2 o O_2) estabiliza la molécula de O_3 recién formada.



El Ozono luego es fotolizado a un átomo de oxígeno y una molécula de oxígeno ante radiación menor a 1180 nm, liberando energía cinética (Wikipedia contributors, 2019f).

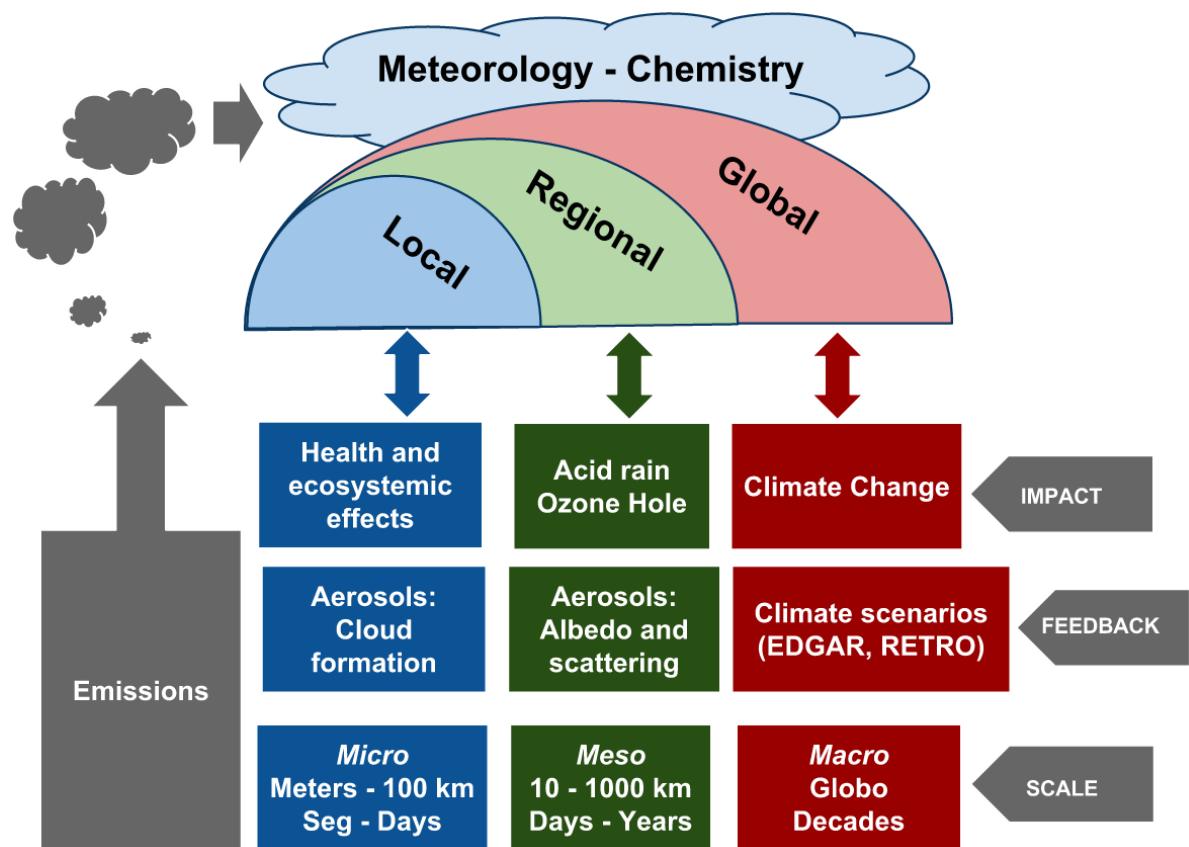
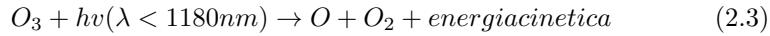
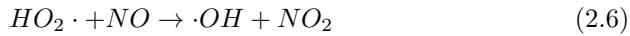
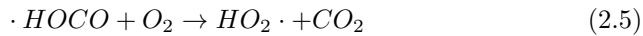


Figure 2.6: Esquema de impacto de emisiones (Ibarra, 2017)

En el caso del ozono troposférico, la reacción comienza cuando el radical hidroxilo $\cdot OH$ oxida el CO generando el radical $\cdot HOCO$ que es inestable y reacciona rápidamente con O_2 generando el peroxi radical $\cdot HO_2$ y CO_2 . El peroxi radical reacciona oxidando NO generando NO_2 y un radical hidroxilo. Luego el NO_2 es fotolizado liberando un átomo de oxígeno que reacciona con la molécula de O_2 generando O_3 .



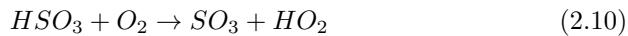
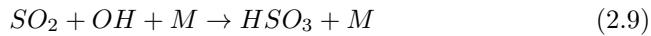
En el caso del ozono troposférico, los ingredientes principales son radiación solar, NO_x y compuestos orgánicos volátiles (Wikipedia contributors, 2019g).

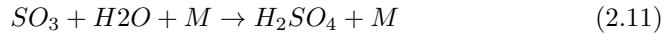


Los mecanismos mostrados son un resumen de complejas reacciones en la atmósfera donde muchas otras reacciones y compuestos juegan un rol.

2.2.2 Especies de azufre

El es un compuesto que reacciona generando importantes contaminantes como sulfatos y ácido sulfúrico, importantes en la lluvia ácida (depósito ácido). Los combustibles tienen azufre y durante la combustión, el azufre es oxidado generando SO_2 , el diesel tiene mayor cantidad de azufre que la gasolina, por lo tanto generando más SO_2 . También existe el dimetilsulfuro (DMS) $(CH_3)_2S$, que es biológico, así como el carbonilo sulfuro (COS), que tiene un largo tiempo de vida que permite su transporte a la estratosfera. A continuación una resumen de los mecanismos de generación de ácido sulfúrico (Brasseur and Jacob, 2017). SO_2 es oxidado por OH





2.2.3 Aerosoles

Son partículas suspendidas desde $\sim 0.001 \mu m$ hasta $100 \mu m$ (luster molecular a gota). Conocido como material particulado (MP), sus caracterización es realizada mayormente en base a su diámetro aerodinámico como se muestra en la siguiente figura (Wikipedia contributors, 2019a)(Brasseur and Jacob, 2017). Existen tres agrupaciones que son:

1. Modo Aiken: núcleos de condensación nuevos (fresh) que condensan (gas) o coagulan (líquidos). Diámetro hasta 100 nm.
2. Modo acumulación: Diámetro entre 100 y 1000 nm.
3. Modo coarse: Diámetro mayor que 1000 nm.

Los aerosoles son clasificados por números, área de superficie y volumen. Normalmente las agencias de medio ambiente miden MP_{10} (diámetro menor \leq que $10 \mu m$) y $MP_{2.5}$ (diámetro menor \leq que $2.5 \mu m$) en $\mu g \cdot m^{-3}$. A continuación una figura mostrando material particulado en Chile, Brasil y China.

Chile *

China

ingrese aquí: <https://aqicn.org/forecast/beijing/>

Osasco, Brasil

NASA WORLDVIEW es un excelente recurso para monitorear aerosoles, incendios y mucho más.

Ingrese a <https://worldview.earthdata.nasa.gov/> y busque aerosol

2.3 Forzantes climáticos y gases de efecto invernadero

El efecto invernadero es un fenómeno que se produce naturalmente en la tierra, sin intervención humana. Por ejemplo, concentraciones biogénicas de CO₂ o el vapor de agua. Sin embargo, el término forzante climático se refiere a estos compuestos químicos que tienen una incidencia en el clima, pues alteran el cambio en irradiancia neta $W \cdot m^{-2}$.

El forzante radiativo es utilizado como predictor de cambio en la media global de temperatura. IPCC (Schimel et al., 1996) define forzamiento radiativo como:

“The radiative forcing of the surface-troposphere system due to the perturbation in or the introduction of an agent (say, a change in greenhouse gas concentrations) is the change in net (down minus up) irradiance (solar plus long-wave; in W m⁻²) at the tropopause AFTER allowing for stratospheric temperatures to

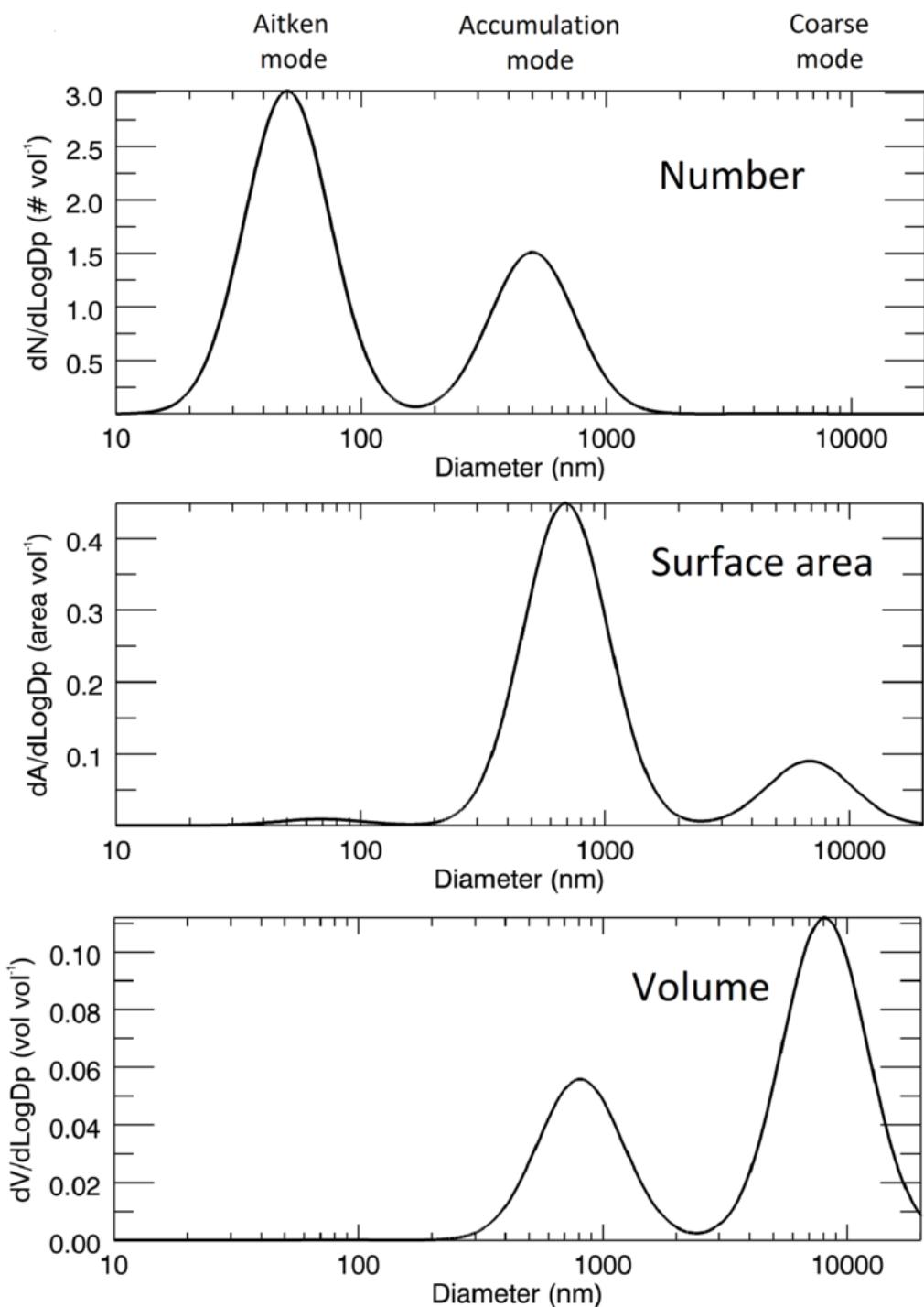


Figure 2.7: Distribucion normalizada (suma es 1000) de aerosoles por diametro aerodinamico, Dominio Publico

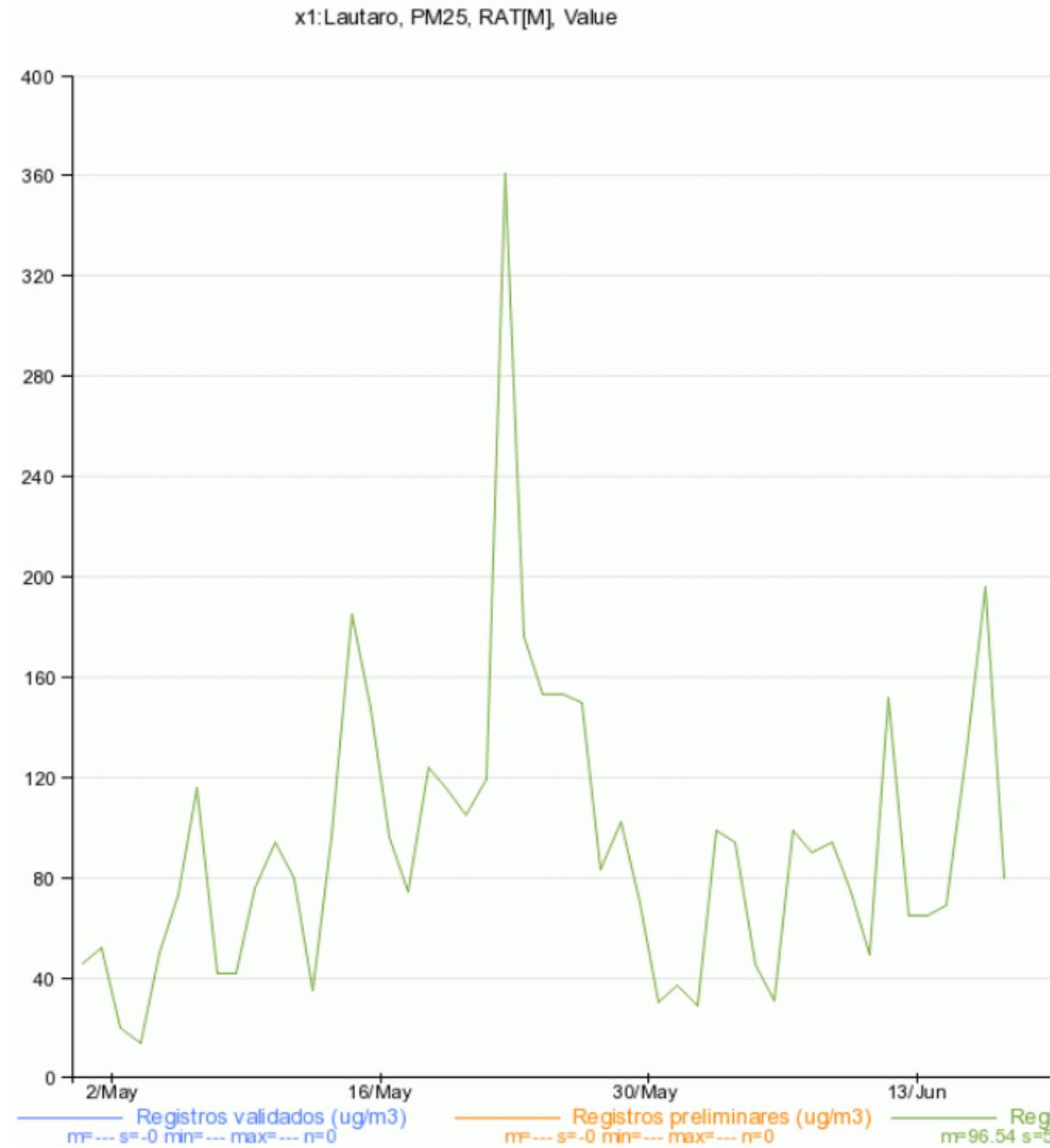


Figure 2.8: MP2.5 $\mu\text{g}/\text{m}^3$ em Lautaro, Chile
[\(https://sinca.mma.gob.cl/index.php/estacion/index/key/870\)](https://sinca.mma.gob.cl/index.php/estacion/index/key/870)

2.3. FORZANTES CLIMATICOS Y GASES DE EFECTO INVERNADERO25

Monitoring points	AQI	Air quality index category	Primary pollutant	PM2.5 fine particles	PM10 respirable particulate matter	CO carbon monoxide	NO2 nitrogen dioxide	O3 ozone 1 hour average	O3 ozone 8 hour average	SO2 sulfur dioxide
Regional station	500	Severe pollution	Particulate matter (PM10)	254	724	0.7	11	76	62	15
Gujiang Baga Township	393	Severe pollution	Particulate matter (PM10)	143	494	0.7	7	78	64	14

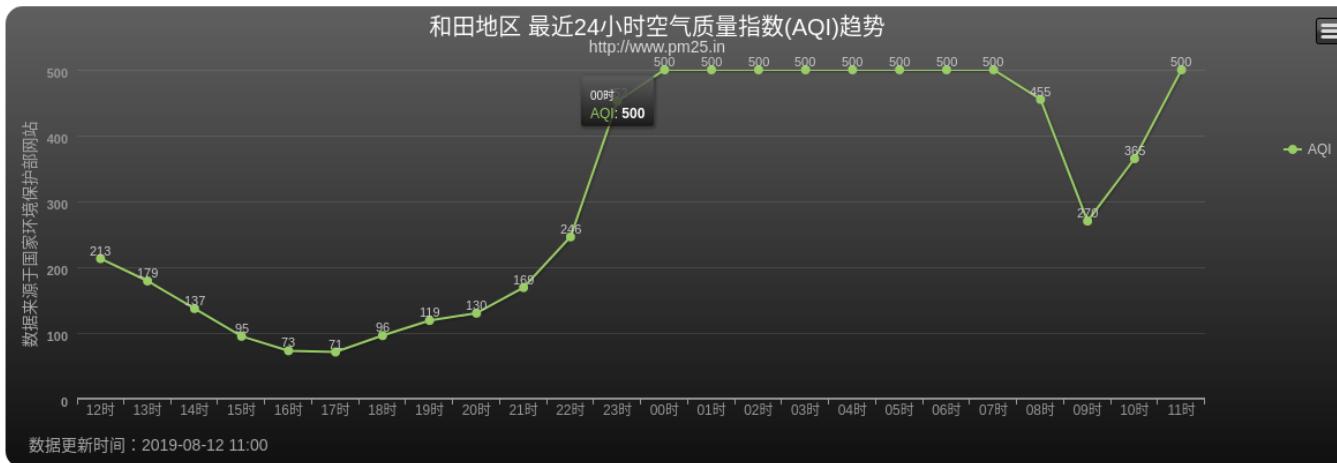
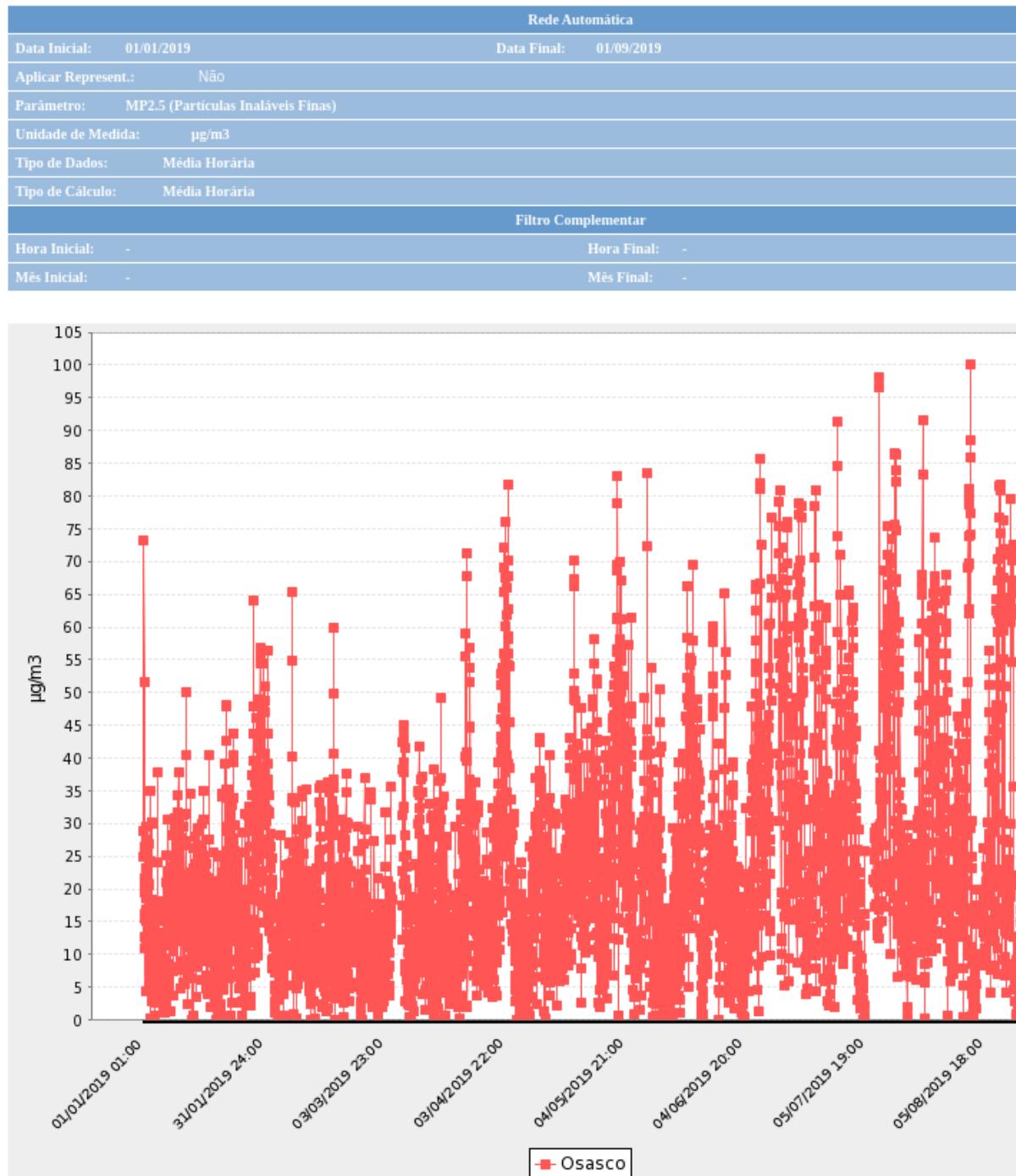


Figure 2.9: MP2.5 ug/m³ em Hotan, China (<http://pm25.in/hetiandi>)

26 CHAPTER 2. IMPACTO DE LAS EMISIONES ANTROPOGÉNICAS EN LA CALIDAD DEL AIRE

Figure 2.10: MP2.5 $\mu\text{g}/\text{m}^3$ Osasco, Brasil (<https://qualar.cetesb.sp.gov.br>)

2.3. FORZANTES CLIMATICOS Y GASES DE EFECTO INVERNADERO27

readjust to radiative equilibrium, but with surface and tropospheric temperatures and state held fixed at the unperturbed values” (Schimel et al., 1996)

Del sitio web de Copernicus (<https://atmosphere.copernicus.eu/climate-forcing>:

“Climate forcing measures the imbalance in the Earth’s energy budget caused by a perturbation of the climate system, for example changes in atmospheric composition driven by human activities. **Climate forcing, also known as Radiative Forcing, therefore determines the change in globally-averaged temperature change due to the natural or human-induced changes to the energy budget.** Increases in greenhouse gas concentrations over the industrial era are responsible for a positive climate forcing, causing a gain of energy in the climate system. In contrast, changes in atmospheric aerosol concentrations result in a negative climate forcing leading to a loss of energy. It is the balance between these various climate forcings that drive the change in global temperature.” <https://atmosphere.copernicus.eu/climate-forcing>

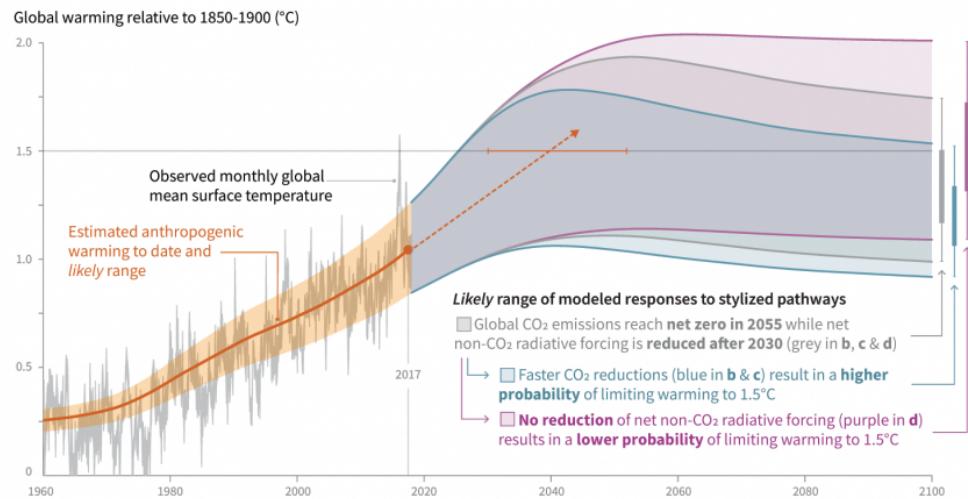
Esta figura muestra un resumen de los compuestos y su forzante radiativo (FR) (Stocker, 2014)

2.3.1 Efectos directos e indirectos de aerosoles

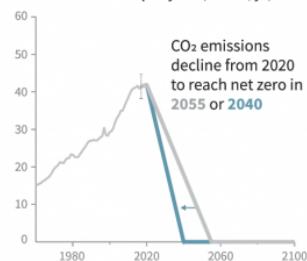
- Directo: Influencia en balance radiativo.
- Indirecto: Influencia en el clima.
- Primer efecto indirecto: microfisica de la precipitacion.
- Segundo efecto indirecto: Cantidad de lluvia.

Cumulative emissions of CO₂ and future non-CO₂ radiative forcing determine the probability of limiting warming to 1.5°C

a) Observed global temperature change and modeled responses to stylized anthropogenic emission and forcing pathways

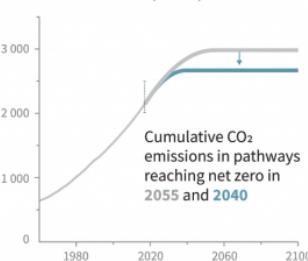


b) Stylized net global CO₂ emission pathways
Billion tonnes CO₂ per year (GtCO₂/yr)



Faster immediate CO₂ emission reductions limit cumulative CO₂ emissions shown in panel (c).

c) Cumulative net CO₂ emissions
Billion tonnes CO₂ (GtCO₂)



Maximum temperature rise is determined by cumulative net CO₂ emissions and net non-CO₂ radiative forcing due to methane, nitrous oxide, aerosols and other anthropogenic forcing agents.

d) Non-CO₂ radiative forcing pathways
Watts per square metre (W/m²)

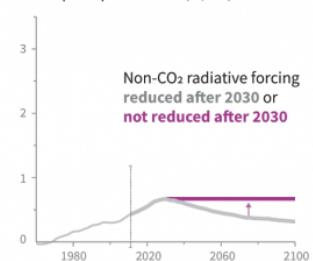


Figure 2.11: Cambio en la temperatura media debido a forzantes radiativas (<https://atmosphere.copernicus.eu/climate-forcing>)

2.3. FORZANTES CLIMATICOS Y GASES DE EFECTO INVERNADERO29

Contributions to observed surface temperature change over the period 1951–2010

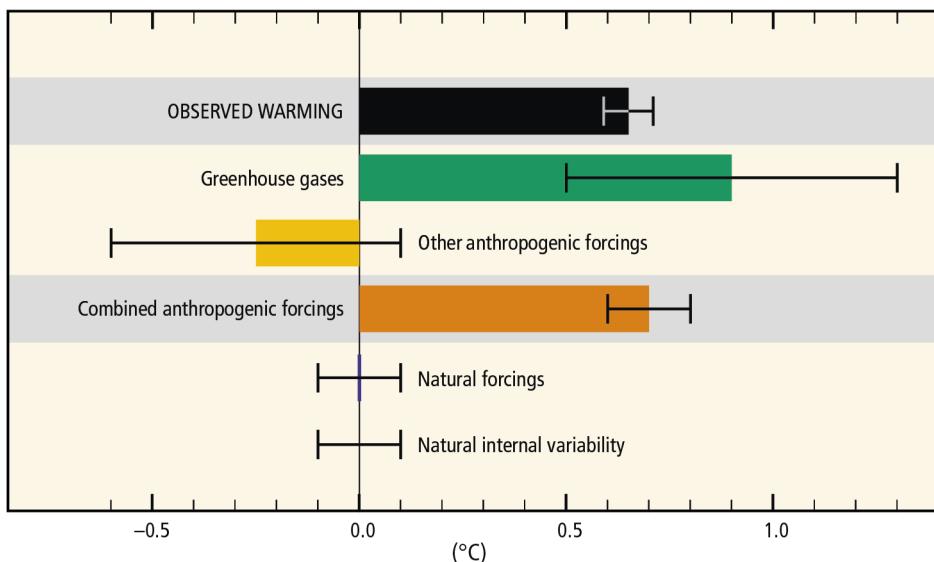
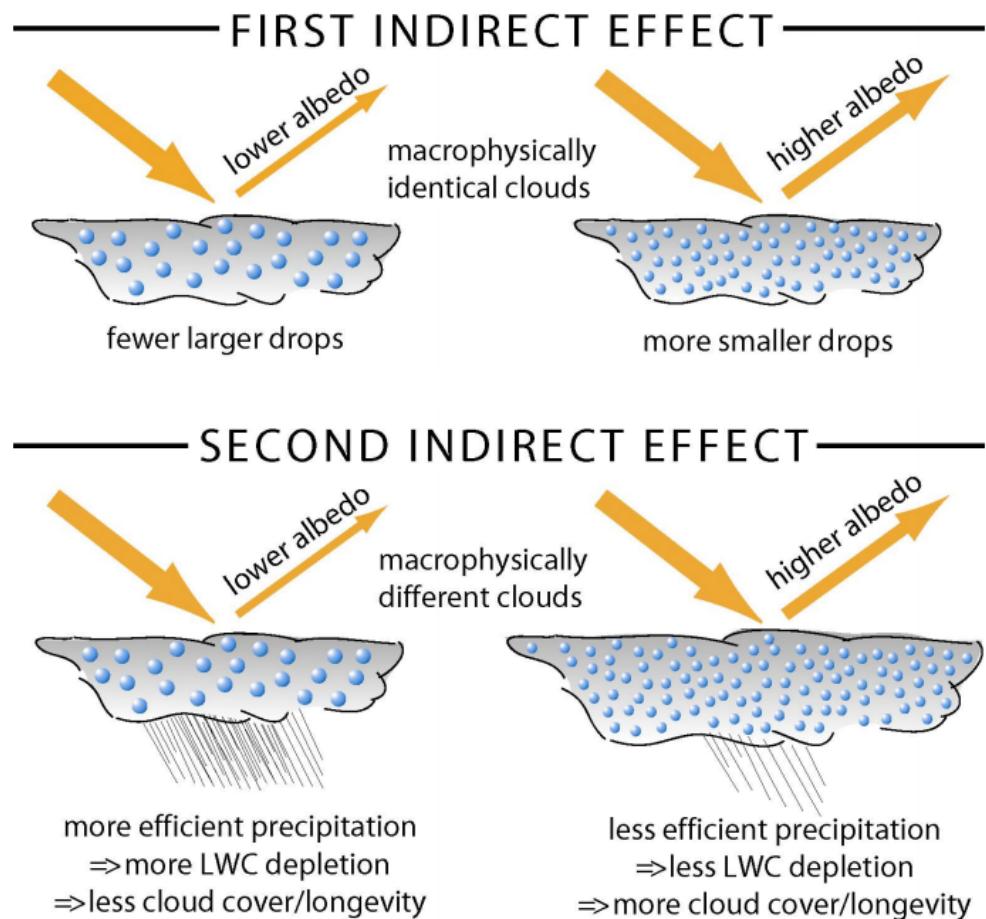


Figure 2.12: Cambio en la temperatura media debido a forzantes radiativas
<https://atmosphere.copernicus.eu/climate-forcing>



\begin{figure}

30CHAPTER 2. IMPACTO DE LAS EMISIONES ANTROPOGÉNICAS EN LA CALIDAD DEL AIRE

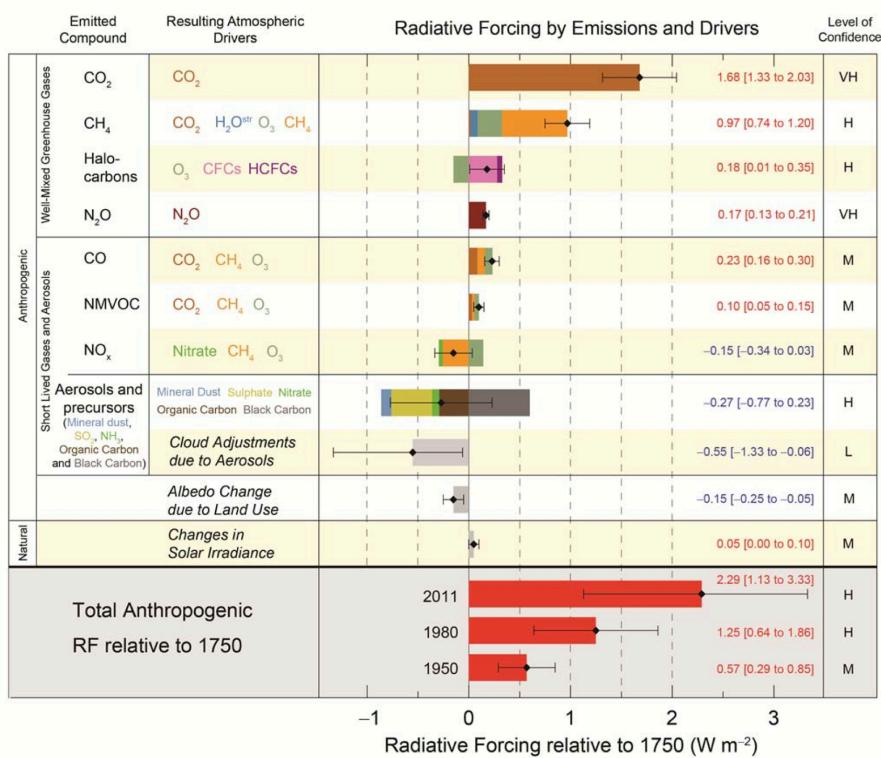


Figure 2.13: Forzante radiativo (https://atmosphere.copernicus.eu/climate-forcing)

2.3. FORZANTES CLIMATICOS Y GASES DE EFECTO INVERNADERO31

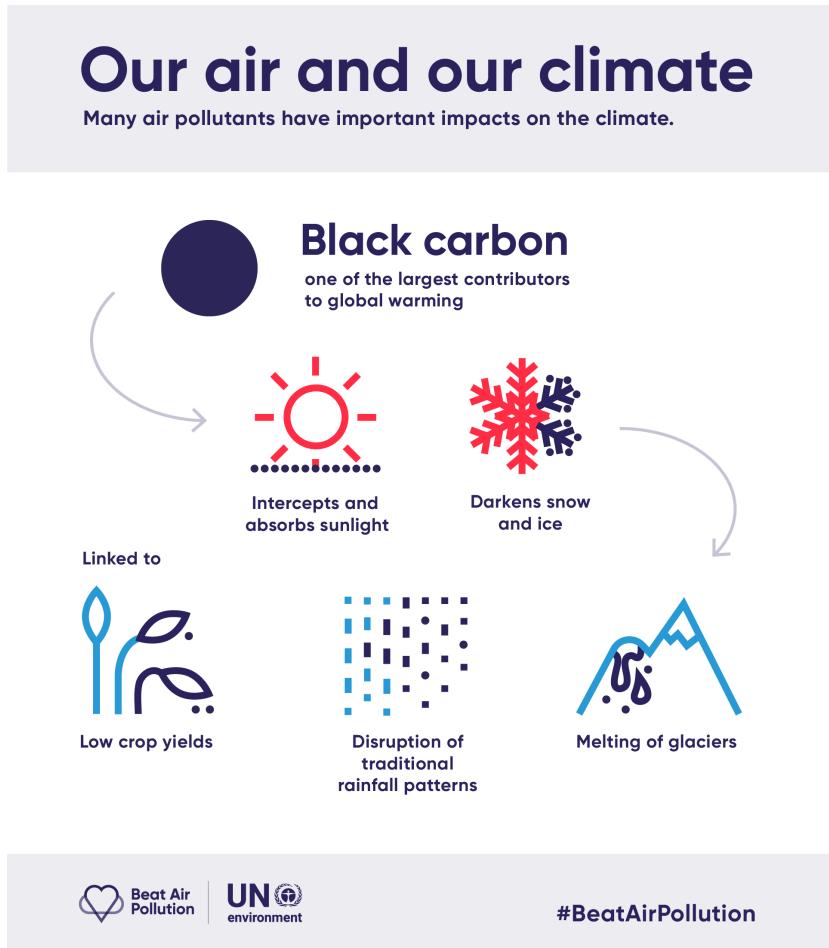


Figure 2.14: UNEP (<https://twitter.com/UNEnvironment/status/1150404638501416961/photo/1>

\caption{Forzantes indirectos (Robwood, 2019) (https://atmos.uw.edu/~robwood/teaching/591/ATMS_591_Albrecht_1989.pdf)} \end{figure}

2.3.2 Black Carbon - Carbono negro

Son partículas producto de la combustión incompleta de combustibles con origen antropogénico y biológico. Considerado el segundo gas de efecto invernadero después de CO₂ (Bond et al., 2013).



Figure 2.15: Fuentes de emisiones

2.4 Emisiones y sus fuentes

Un inventario de emisiones es la compilacion del flujo de masa de contaminantes emitida en un lugar y tiempo determinado.

Son definidos por la ecuacion basica de :

$$E = FE \cdot NA \quad (2.12)$$

Donde E es la emision, FE es el factor de emision y NA el nivel de actividad. Por ejemplo, si queremos saber las emisiones de vehiculos, FE esta en **g/km**, NA es la cantidad de vehiculos veces la distancia que recorren **km** en un tiempo determinado, luego E es la masa **g** de contaminantes.

2.4.1 Tipos

Ya que ya sabemos los contaminantes y su forzamiento radiativo, podemos ver cuales son las fuentes de ellos, desde el punto de vista de contaminacion atmosferica y clima. Existen varios tipos de fuentes de emisiones:

Pueden tener aplicaciones de gestion y cientificas:

Type of application	Inventory examples
Policy applications	National greenhouse gas inventories, submitted to the United Nations Framework Convention on Climate Change (UNFCCC [10])
	National air pollutant inventories submitted to the United Nations Economic Commission for Europe's Convention on Long Range Transboundary air Pollution (LRTAP [11])
	Pollutant Release and Transfer Registers (PRTR) like the USA Toxics Release Inventory (TRI, [12]) and the European Union's EPER system (EPER [13])
Scientific applications	The GEIA emission databases (GEIA [14])
	The EDGAR emission inventories (EDGAR [15])
	The TNO Emissions Assessment Model (TEAM [16, 17])

Figure 2.16: Tipos de inventarios (Pules and Helsinga, 2013)

Por ejemplo, IPCC entrega guías para que los países compilen inventarios de emisiones de efecto invernadero:

Ahora podemos ver como ha sido la serie de emisiones globales de gases de efecto invernadero:

2.4.2 Calidad y criterio

Un inventario de emisiones debe considerar la aseguración de la calidad. En este sentido, los científicos buscan errores, puntos débiles. Los tomadores de decisiones (UNFCCC-IPCC) buscan consenso. Por lo tanto, la calidad para los inventarios científicos consiste en que las estimaciones tienen que ser confirmadas, calibradas y validadas. Y en el plano político, un inventario de calidad genera acuerdo y protocolo que todos deben atacar (VEA CLRTAP)

2.4.3 Dimensiones

Como **científicos** necesitamos responder las siguientes preguntas:

- Que: Que contaminante (específico) es emitido.
- Como: Cual es el proceso.
- Cuando: Caracterización temporal.
- Onde: Caracterización temporal.

2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories

2019 Refinement has been accepted!

The following files are the advance version of the Overview Chapter and the underlying "2019 Refinement to the 2006 Guidelines for National Greenhouse Gas Inventories" including the Glossary as adopted/accepted on 12 May 2019 at the 49th session of the IPCC [<Decision - IPCC-XLIX-9 - Adoption and Acceptance of 2019 refinement>](#). They are subject to final copy-edit and layout prior to its final publication.

Added is the Note with the explanation of the technical corrections of editorial nature made to the accepted 2019 Refinement.

More information on the development of the 2019 Refinement can be found [here](#).

Overview Chapter

Volume 1 General Guidance and Reporting

Volume 2 Energy

Volume 3 Industrial Processes and Product Use

Volume 4 Agriculture, Forestry and Other Land Use

Volume 5 Waste

Figure 2.17: Actualización en las guías de inventarios de emisiones IPCC (<https://www.ipcc-nppg.iges.or.jp/public/2019rf/index.html>)

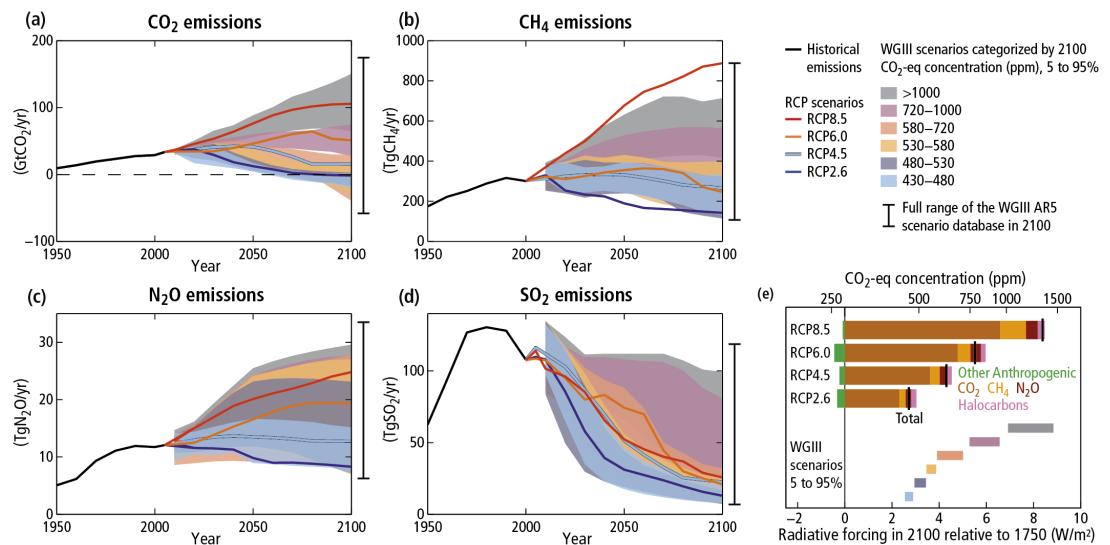
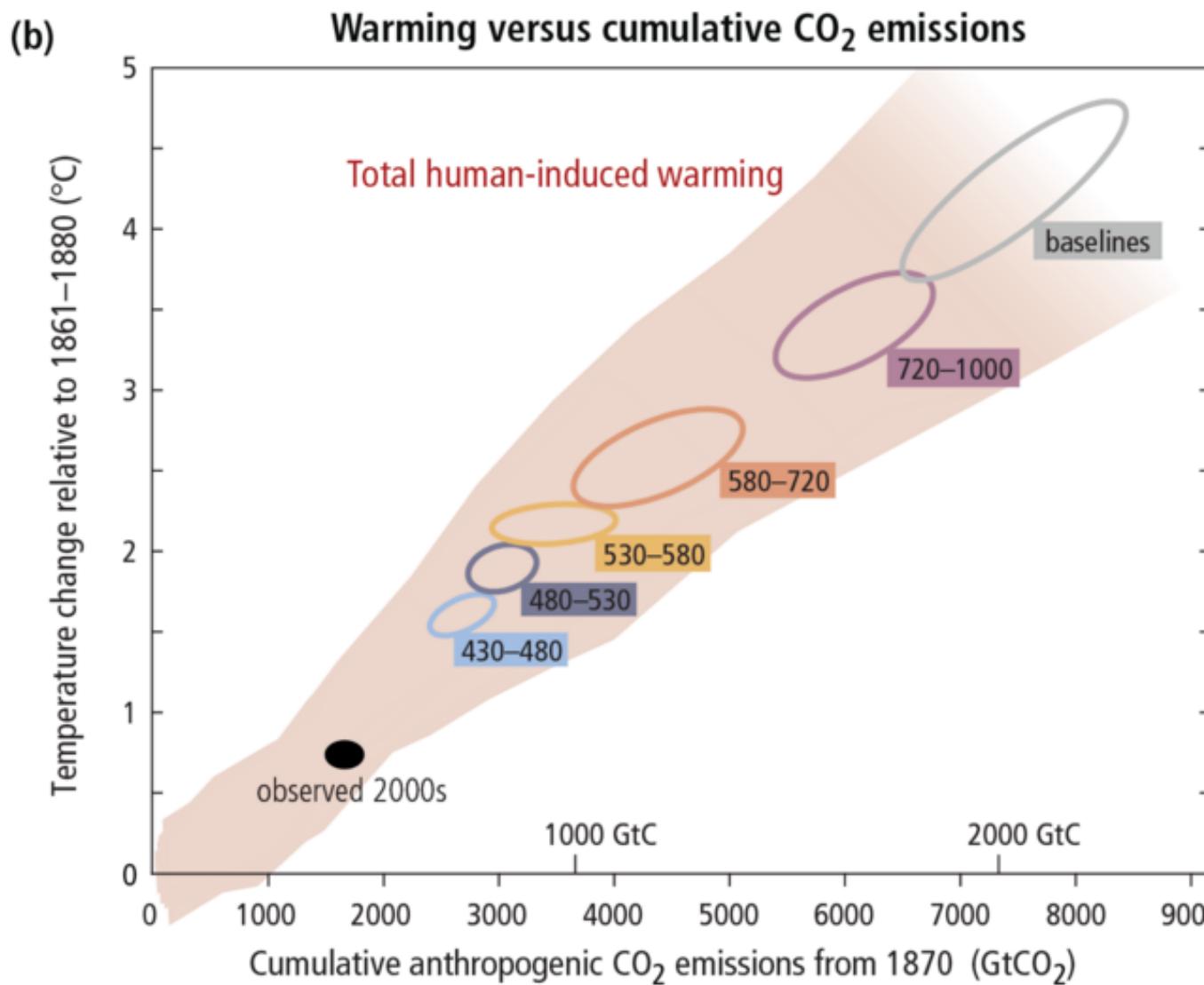


Figure 2.18: Serie de emisiones)

Figure 2.19: CO₂ y temperatura (<http://www.ipcc.ch/report/graphics/>)

** Mapas interactivos de inventario de emisiones vehiculares de Sao Paulo (Ibarra 2017)**

- PM2.5
- NOx
- CO
- HC

2.4.4 Categorías llave

En los inventarios de emisiones muchas veces pocas categorías emiten la mayor cantidad de contaminantes. Estas son las **categorías llave** y es en ellas donde debemos invertir la mayor cantidad de esfuerzo y recurso económico.

2.4.5 Bottom-up y top-down

En las guías de emisiones europeas, (2016) muestra la definición de los enfoques bottom-up y top-down.

2.4.6 Ejemplo de factores de emisión (real-time GPS)

2.5 Nuevos modelos de emisiones: VEIN

En el último par de años se han publicado nuevos modelos para el cálculo de emisiones, como el Vehicular Emissions Inventory model (Ibarra-Espinosa et al., 2018) que consiste en una librería en R (<https://atmoschem.github.io/vein/>).

2.6 Nuevos datos, satélites y modelación inversa

La modelación inversa está siendo utilizada para corregir y estimar emisiones con los nuevos datos de satélite disponibles (Kaiser et al., 2018):

Columna total de NO₂ en europa con Sentinel 5P librería stars (Pebesma, 2019)

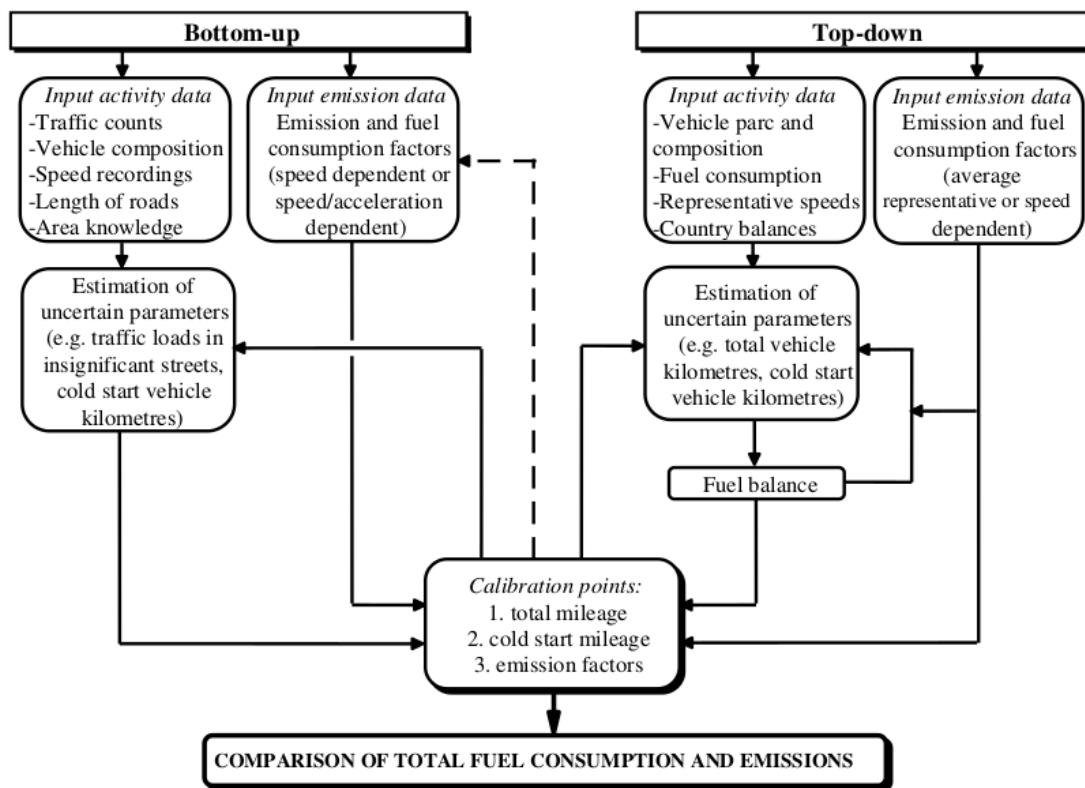


Figure 2.20: Bottom up y top-down

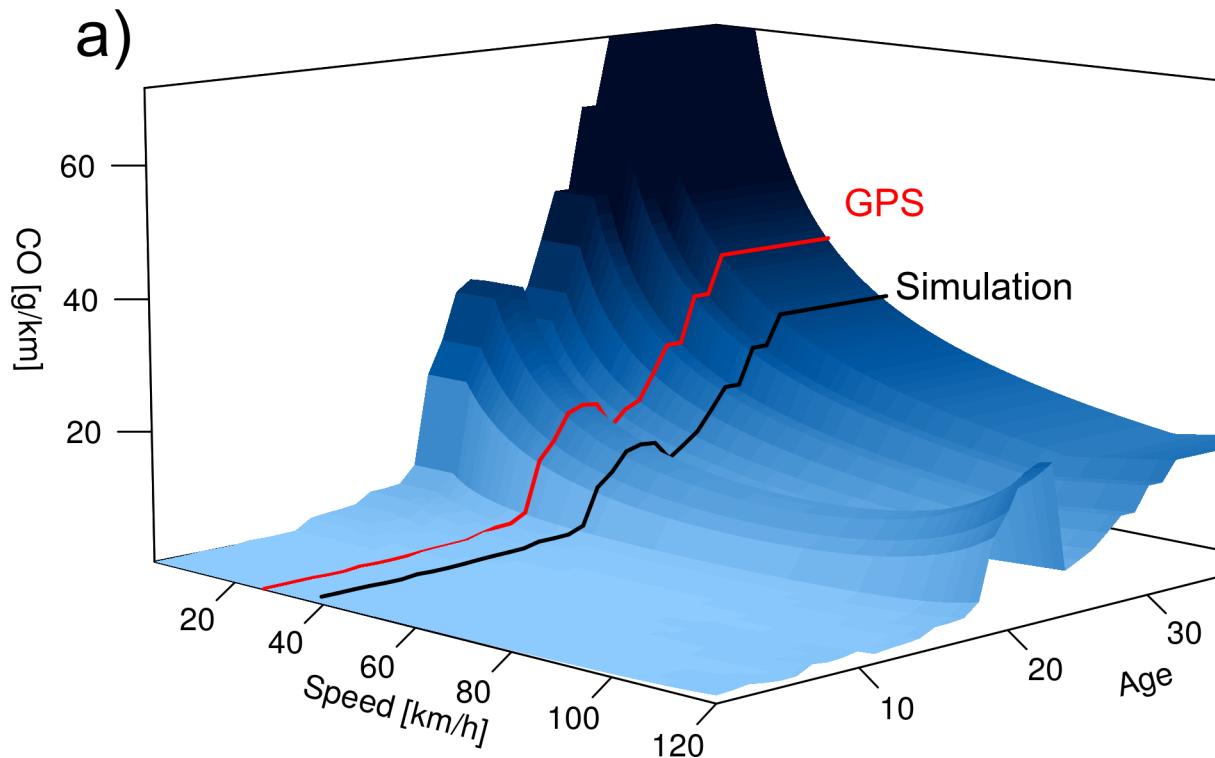


Figure 2.21: (Ibarra-Espinosa et al., 2019 under-review)

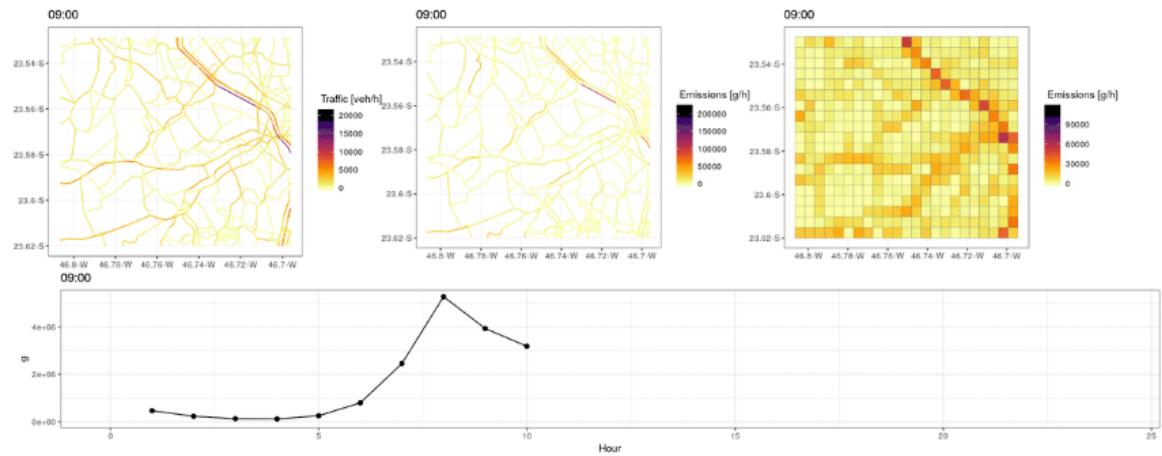
Figure 2.22: (Ibarra-Espinosa et al., 2018, <https://atmoschem.github.io/vein/>)



Figure 2.23: (Ibarra-Espinosa et al., 2018, <https://atmoschem.github.io/vein/>)

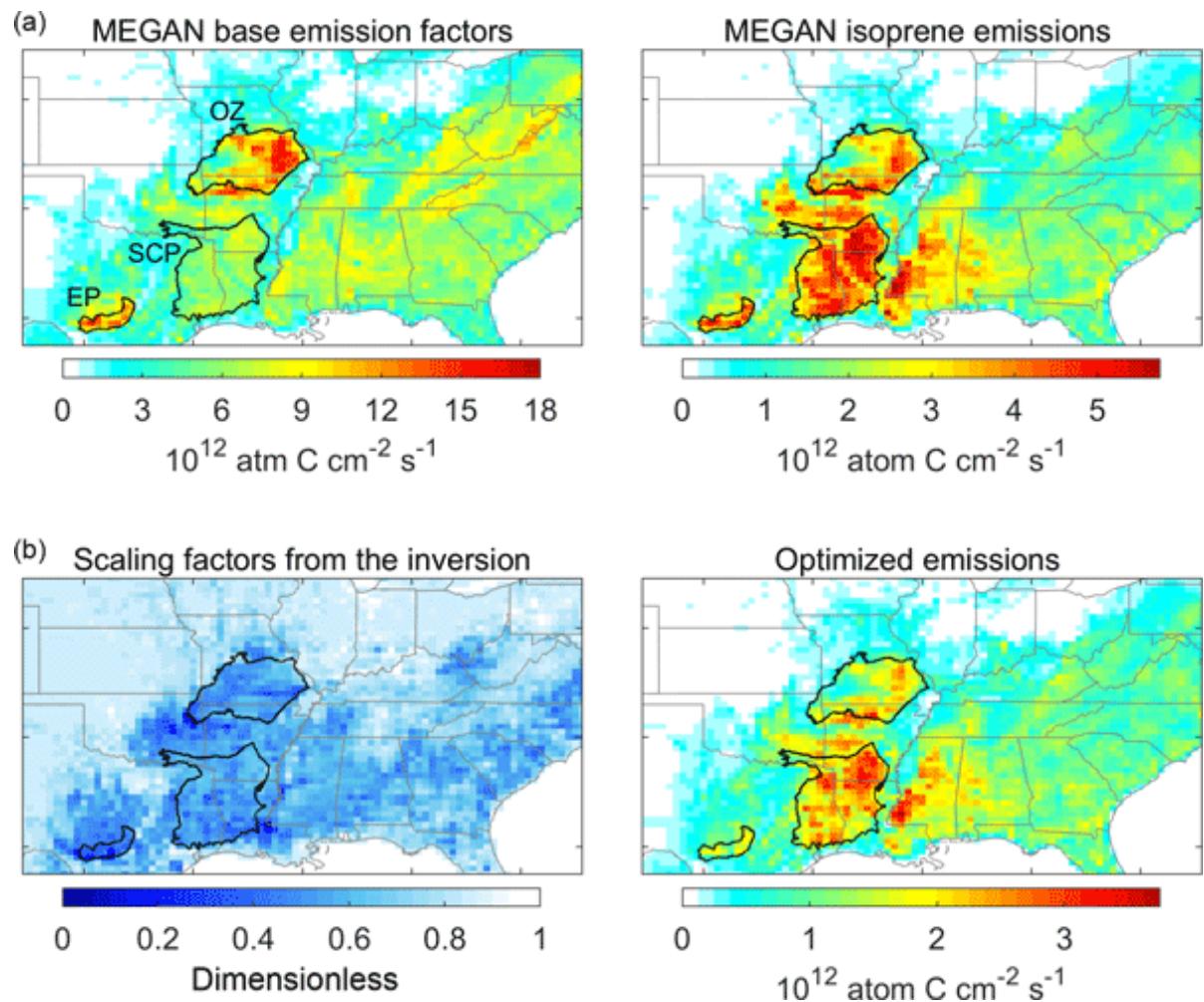


Figure 2.24: (Kaiser et al., 2018)

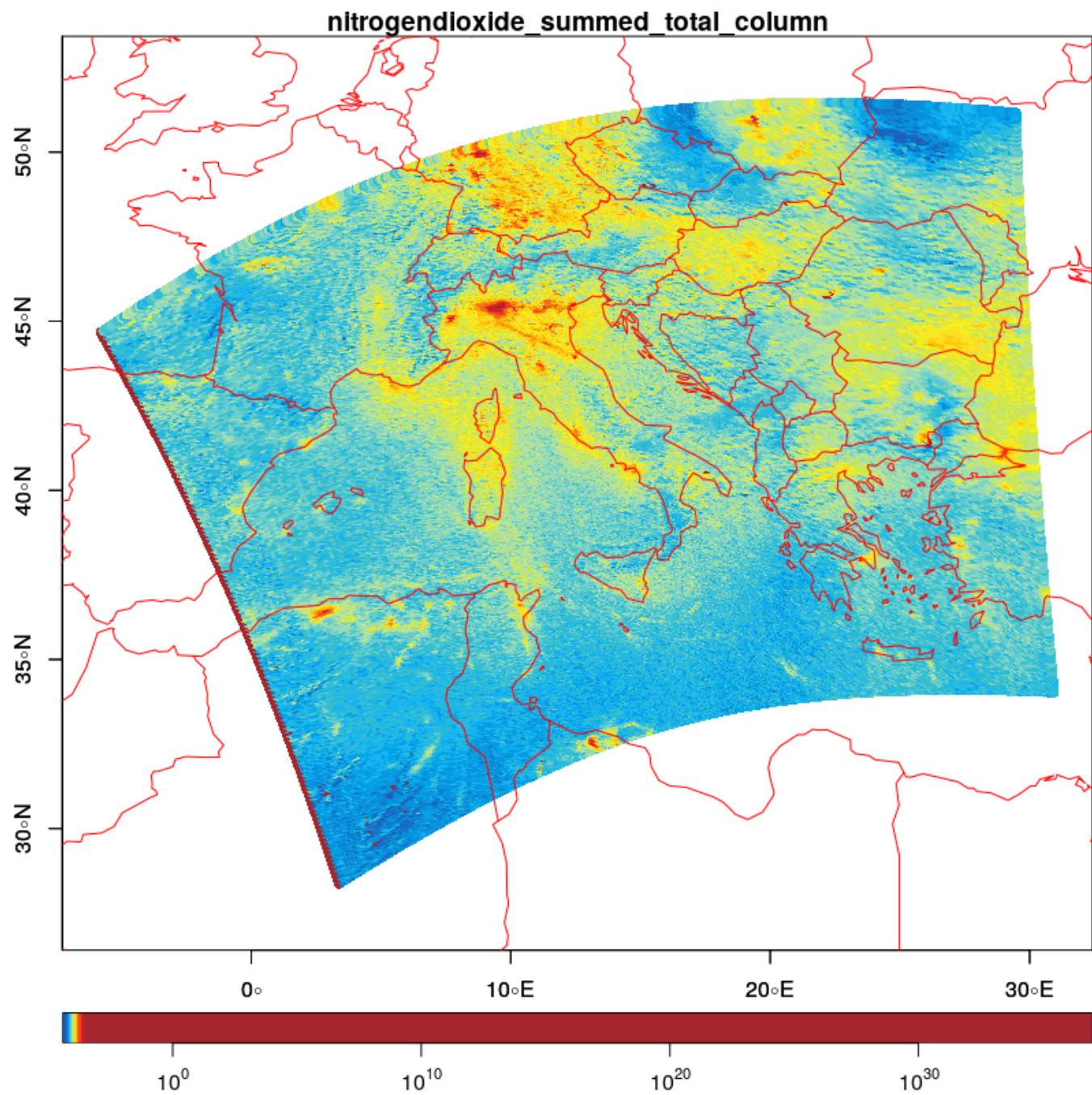


Figure 2.25: STARS! Sentinel 5P

Chapter 3

Taller VECTORES: Aplicación de software de información geográfica y modelado

gracias Dra. Patricia Matus

Nota: Existen diferentes recursos para aprender R, sf, stars y raster.
Use Google, Baidu o duckduckgo

- www.github.com/r-spatial/sf
- www.github.com/r-spatial/stars
- www.github.com/rspatial/raster
- <https://bookdown.org/>
- <https://geocompr.robinlovelace.net/>
- <https://www.datascienceatthecommandline.com/>
- <https://bookdown.org/rdpeng/rprogdatascience/>

NOTA: SI ALGUNA VEZ TE PIERDES USANDO ALGUNA FUNCION DE R, ESCRIBE ?FUNCION Por ejemplo, si tenemos dudas con la función aggregate, plot, read.csv

```
?aggregate  
?plot  
?read.csv
```

NOTA: PARA COMENTAR, ESCRIBA #

Por ejemplo,

?aggregate #importante! Esto abre la pagina del manual de la función aggregate

3.1 Instalacion

Por favor instale las siguientes librerías en R, copie y pegue (cole) en R:

```
install.packages(
  c("sf", "stars", "cptcity", "ggplot2", "raster", "ncdf4",
    "RNetCDF", "maps", "data.table", "readxl")
)
```

Para correr el ejemplo de stars con Sentinel 5P instale la siguiente librería (1 Gb)

```
install.packages(
  c("starsdata")
)
```

Para realmente comenzar es necesario mencionar que R es un lenguaje de programación estadístico libre (gratis y abierto) orientado a objeto escrito en C, con bindings directos para C++, C y Fortran.

3.2 R desde Excel, libreoffice, archivos de texto, etc.

A veces debemos obtener datos en hojas de cálculo (LibreOffice Calc o Microsoft Office Excel). Cuando este archivo es pequeño no hay problema, pero la realidad es que cada vez es más frecuente contar con grandes bases de datos y trabajar con un programa de interfaz gráfica se hace difícil. Una de las razones es que la interfaz gráfica consume muchos recursos computacionales, que podrían ser usados para el procesamiento de información. Por lo tanto

3.2.1 El Niño y la Niña

“El Niño (La Niña) is a phenomenon in the equatorial Pacific Ocean characterized by a five consecutive 3-month running mean of sea surface temperature (SST) anomalies in the Niño 3.4 region that is above (below) the threshold of +0.5°C (-0.5°C). This standard of measure is known as the Oceanic Niño Index (ONI).” NOAA

Vamos a leer la temperatura de la superficie del mar para calcular el índice ONI

do (5 min) - Abra <https://www.ncdc.noaa.gov/teleconnections/nao/data.csv>
 - ¿Cómo son los datos? - ¿Cuál es la separación de las columnas?

3.2.2 Usando base

```

url <- "http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/detrend.nino34.ascii."
# Leyendo datos
oni = read.table(url, h = T)
# construyendo la variable de fecha / data
oni$Date <- as.Date(paste0(oni$YR, ifelse(nchar(oni$MON) <
                                2, paste0(0, oni$MON), oni$MON), "01"), format = "%Y-%m-%d")
# Extrayendo el año
oni$Year <- as.integer(strftime(oni$Date, "%Y"))
# Extrayendo el mes
oni$Month <- as.integer(strftime(oni$Date, "%m"))
# Definiendo la función de media móvil
ma <- function(x, n = 3) {
  stats::filter(x, rep(1/n, n), sides = 2)
}
# Calculando la media móvil
oni$ONI <- ma(oni$ANOM, 3)
# Seleccionando las columnas de interés
oni <- oni[, c("Date", "Year", "Month", "ONI")]
# Mostrando las primeras 6 observaciones de cada columna
head(oni)

##           Date Year Month      ONI
## 1 1950-01-01 1950     1       NA
## 2 1950-02-01 1950     2 -1.336667
## 3 1950-03-01 1950     3 -1.166667
## 4 1950-04-01 1950     4 -1.183333
## 5 1950-05-01 1950     5 -1.073333
## 6 1950-06-01 1950     6 -0.850000

```

3.2.3 Usando data.table (mi favorito junto con sf)

data.table es mas rapido que python, julia y spark

<https://h2oai.github.io/db-benchmark/>

```

library(data.table)
# Leyendo datos
oni = fread(url, h = T)
# construyendo la variable de fecha / data
oni$Date <- as.Date(paste0(oni$YR, ifelse(nchar(oni$MON) <
                                2, paste0(0, oni$MON), oni$MON), "01"), format = "%Y-%m-%d")
# Extrayendo el año
oni$Year <- as.integer(strftime(oni$Date, "%Y"))
# Extrayendo el mes
oni$Month <- as.integer(strftime(oni$Date, "%m"))

```

```
# Definiendo la funcion de media movil
ma <- function(x, n = 3) {
  stats::filter(x, rep(1/n, n), sides = 2)
}
# Calculando la media movil
oni$ONI <- ma(oni$ANOM, 3)
# Seleccionando las columnas de interes
oni <- oni[, c("Date", "Year", "Month", "ONI")]
# mostrando las primeras 6 observaciones de cada columna
oni
class(oni)
```

Otras librerías para leer datos tabulares (planihas - spreadsheets):

- readxl
- readr

3.2.4 summary

```
summary(oni)
```

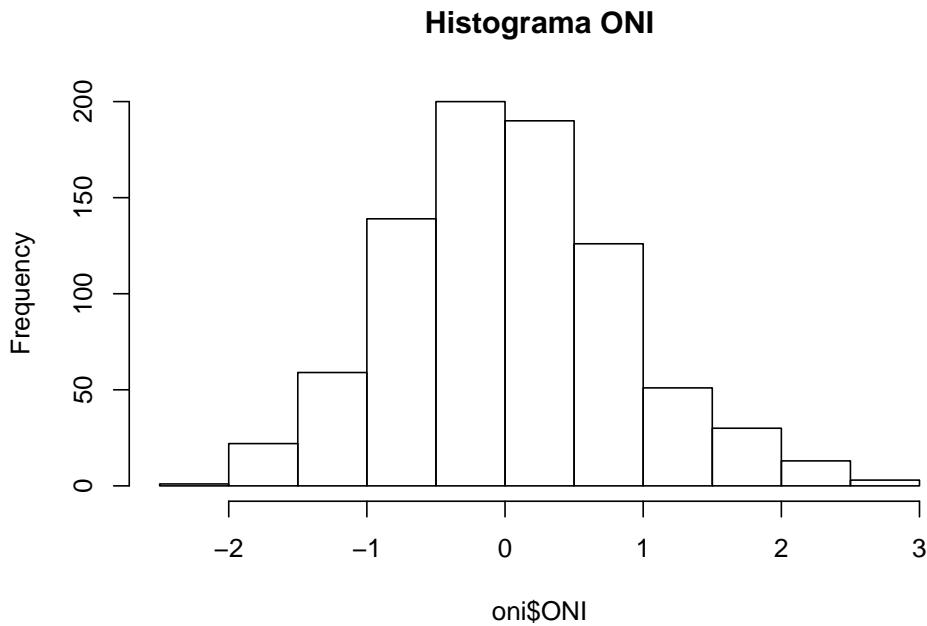
	Date	Year	Month	ONI
## Min.	:1950-01-01	Min. :1950	Min. : 1.000	Min. :-2.033333
## 1st Qu.	:1967-05-24	1st Qu.:1967	1st Qu.: 3.000	1st Qu.:-0.536667
## Median	:1984-10-16	Median :1984	Median : 6.000	Median :-0.003333
## Mean	:1984-10-15	Mean :1984	Mean : 6.481	Mean : 0.032758
## 3rd Qu.	:2002-03-08	3rd Qu.:2002	3rd Qu.: 9.000	3rd Qu.: 0.560833
## Max.	:2019-08-01	Max. :2019	Max. :12.000	Max. : 2.643333
##				NA's :2

3.2.5 Ploteando datos con base

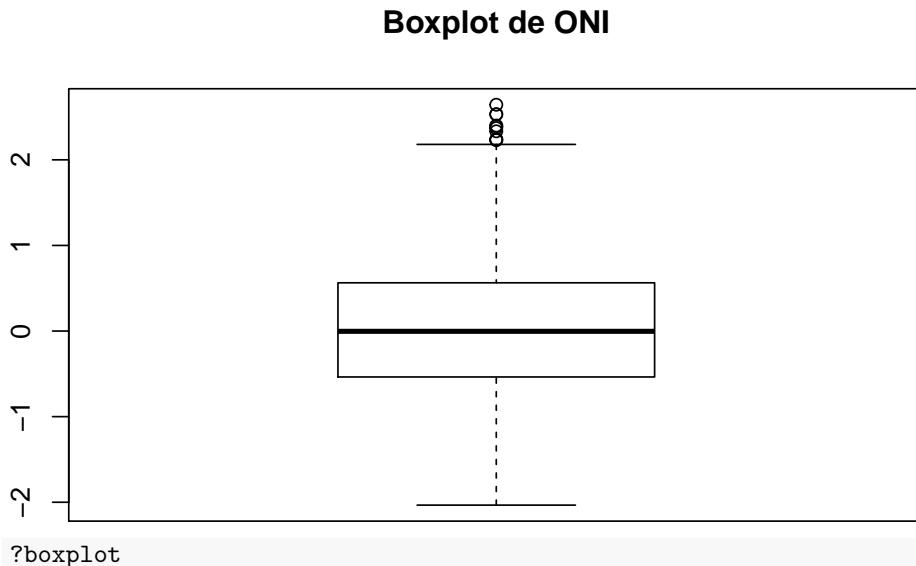
```
?plot
```

```
## Help on topic 'plot' was found in the following packages:
##
##   Package           Library
##   sf                /home/sergio/R/x86_64-pc-linux-gnu-library/3.6
##   stars             /home/sergio/R/x86_64-pc-linux-gnu-library/3.6
##   raster            /home/sergio/R/x86_64-pc-linux-gnu-library/3.6
##   graphics          /usr/lib/R/library
##
## 
## Using the first match ...
```

```
hist(oni$ONI,  
     main = "Histograma ONI")
```



```
boxplot(oni$ONI,  
        main = "Boxplot de ONI")
```



```
?boxplot
```

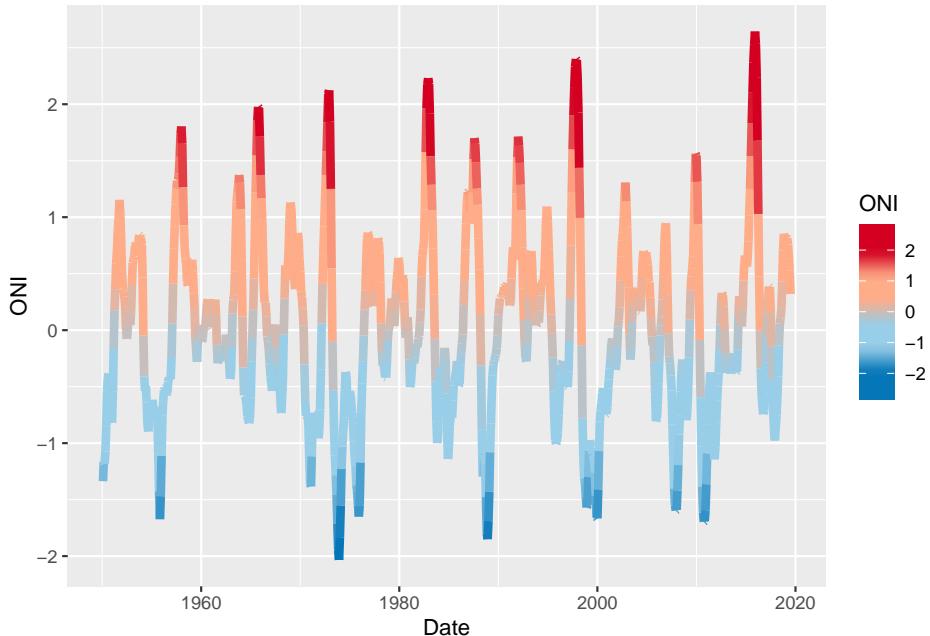
```
## Help on topic 'boxplot' was found in the following packages:
```

```

##                                     Library
##   Package           raster           /home/sergio/R/x86_64-pc-linux-gnu-library/3.6
##   graphics          graphics         /usr/lib/R/library
##
##                                     Using the first match ...
library(ggplot2)
library(cptcity)
ggplot(oni,
       aes(x = Date, y = ONI, colour = ONI)) +
  geom_line(size = 2) +
  scale_colour_gradientn(colours = rev(cpt(find_cpt("cb_div_RdB")[2])),
                         limit = c(-2.7, 2.7))

## Don't know how to automatically pick scale for object of type ts. Defaulting to con-

```



3.3 Vectores

GDAL es la librería de abstracción de datos espaciales. Permite leer diferentes tipos de datos de tipo vectores y [raster](<https://gdal.org/drivers/raster/index.html>). Algunos formatos de vectores:

- ESRI Shapefile
- Google Earth Engine Data API

- GeoJSON para html
- GeoPackage que es el nuevo formato estandard para la representacion de datos espaciales. Ver aqui. Nuevo estandard ISO (puntos, areas, lineas, volume, etc) GeoPackage, creado por la Open Geospatial Consortium, es un formato abierto y no propietario de datos espaciales

sf significa Spatial Features y es una libreria con interface para las librerias GDAL, GEOS y PROJ que te premite trabajar con datos espaciales de una forma muy parecida que con una hoja de calculo normal. Como sf esta enlazada con GDAL, permite leer todos los formatos de vector de GDAL directamente en R.

Vamos a leer datos de Argentina en formato GeoJSON

...

Es demasiado OBVIO que podemos leer datos shapefile .shp, por eso para cambiar un poco vamos a leer GeoJSON

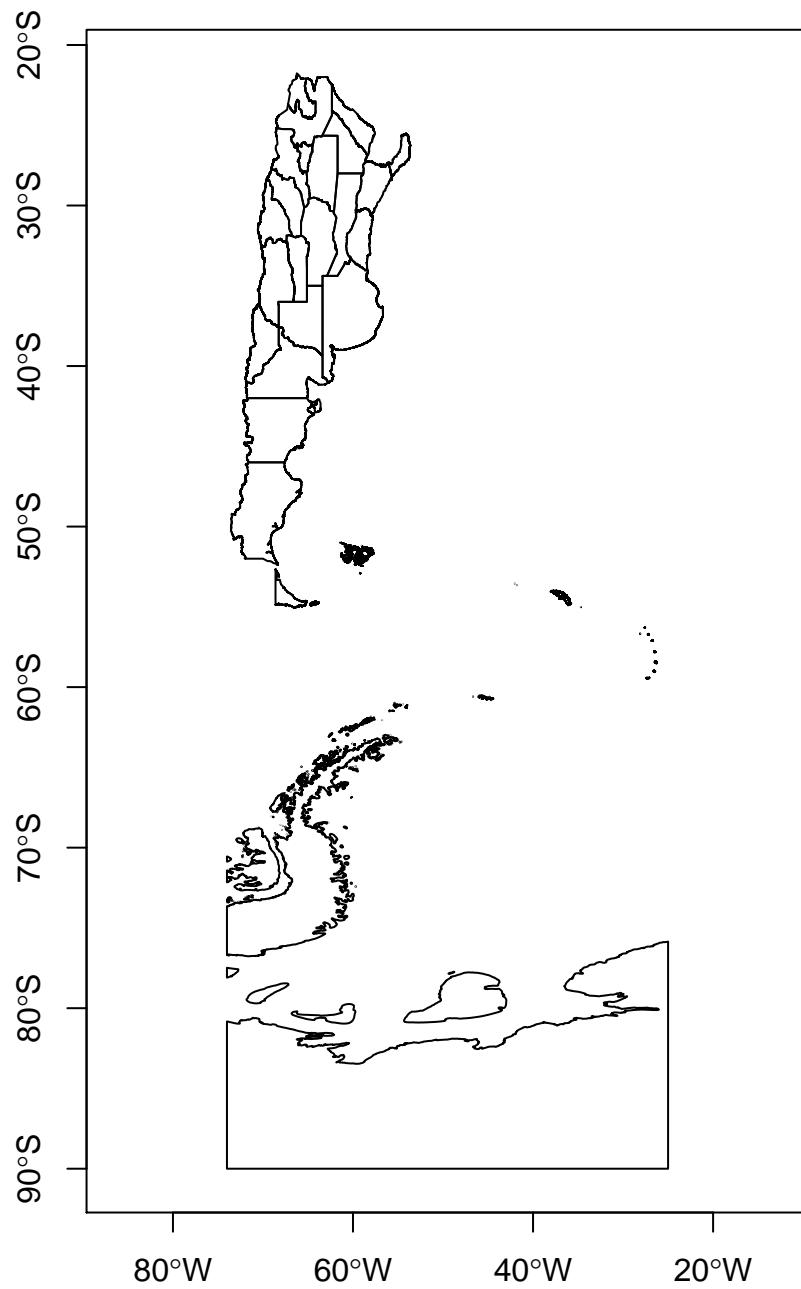
datos bajados de IGN

```
library(sf)
df <- read_sf("dados/provincia.json")
head(df)

## Simple feature collection with 6 features and 8 fields
## geometry type:  MULTIPOLYGON
## dimension:      XYZ
## bbox:            xmin: -71.96569 ymin: -41.10059 xmax: -58.33515 ymax: -29.50042
## epsg (SRID):    4326
## proj4string:    +proj=longlat +datum=WGS84 +no_defs
## # A tibble: 6 x 9
##   entidad objeto fna   gna   nam   in1   fdc   sag
##   <int> <chr> <chr> <chr> <chr> <chr> <chr>
## 1       0 Provi~ Ciud~ Ciud~ Ciud~ 02   Geog~ IGN
## 2       0 Provi~ Prov~ Prov~ Neuq~ 58   Geog~ IGN
## 3       0 Provi~ Prov~ Prov~ La P~ 42   Geog~ IGN
## 4       0 Provi~ Prov~ Prov~ Mend~ 50   Geog~ IGN
## 5       0 Provi~ Prov~ Prov~ San ~ 74   Geog~ IGN
## 6       0 Provi~ Prov~ Prov~ Córd~ 14   Geog~ IGN
## # ... with 1 more variable: geometry <MULTIPOLYGON [°]>
names(df)

## [1] "entidad"   "objeto"     "fna"        "gna"        "nam"        "in1"
## [7] "fdc"        "sag"        "geometry"
plot(df["geometry"],
      main = "Provincias de Argentina",
      axes = TRUE)
```

Provincias de Argentina



Podemos seleccionar columnas y líneas igual a la forma que lo hicimos anteriormente, y ploteamos estos datos.

```
head(df)

## Simple feature collection with 6 features and 8 fields
## geometry type:  MULTIPOLYGON
## dimension:      XYZ
## bbox:            xmin: -71.96569 ymin: -41.10059 xmax: -58.33515 ymax: -29.50042
## epsg (SRID):   4326
## proj4string:    +proj=longlat +datum=WGS84 +no_defs
## # A tibble: 6 x 9
##   entidad objeto fna  gna  nam  in1  fdc  sag
##   <int> <chr> <chr> <chr> <chr> <chr> <chr>
## 1       0 Provi~ Ciud~ Ciud~ 02     Geog~ IGN
## 2       0 Provi~ Prov~ Prov~ Neuq~ 58     Geog~ IGN
## 3       0 Provi~ Prov~ Prov~ La P~ 42     Geog~ IGN
## 4       0 Provi~ Prov~ Prov~ Mend~ 50     Geog~ IGN
## 5       0 Provi~ Prov~ Prov~ San ~ 74     Geog~ IGN
## 6       0 Provi~ Prov~ Prov~ Córd~ 14     Geog~ IGN
## # ... with 1 more variable: geometry <MULTIPOLYGON [°]>
```

Vemos muchos datos, pero vamos a ver los datos unios por departamento. No estoy mostrando aqui pues son 435, pero ya se que vamos a seleccionar “Departamento Belgrano”

```
unique(df$fna)

##  [1] "Ciudad Autónoma de Buenos Aires"
##  [2] "Provincia del Neuquén"
##  [3] "Provincia de La Pampa"
##  [4] "Provincia de Mendoza"
##  [5] "Provincia de San Luis"
##  [6] "Provincia de Córdoba"
##  [7] "Provincia de Santa Fe"
##  [8] "Provincia de Entre Ríos"
##  [9] "Provincia de San Juan"
## [10] "Provincia de La Rioja"
## [11] "Provincia de Catamarca"
## [12] "Provincia de Tucumán"
## [13] "Provincia de Jujuy"
## [14] "Provincia del Chaco"
## [15] "Provincia de Formosa"
## [16] "Provincia de Santiago del Estero"
## [17] "Provincia de Tierra del Fuego, Antártida e Islas del Atlántico Sur"
## [18] "Provincia de Santa Cruz"
## [19] "Provincia del Chubut"
## [20] "Provincia de Río Negro"
## [21] "Provincia de Buenos Aires"
## [22] "Provincia de Corrientes"
```

52CHAPTER 3. TALLER VECTORES: APLICACIÓN DE SOFTWARE DE INFORMACIÓN GEOGRÁFICA

```
## [23] "Provincia de Misiones"  
## [24] "Provincia de Salta"
```

La provincia 17 tiene un nombre muy largo, lo cambiamos asi:

```
df$fna[17] <- "Provincia Antartida"
```

De todas maneras, el campo fna tiene muchas veces la pabra “Provincia”, camos a eliminarla.

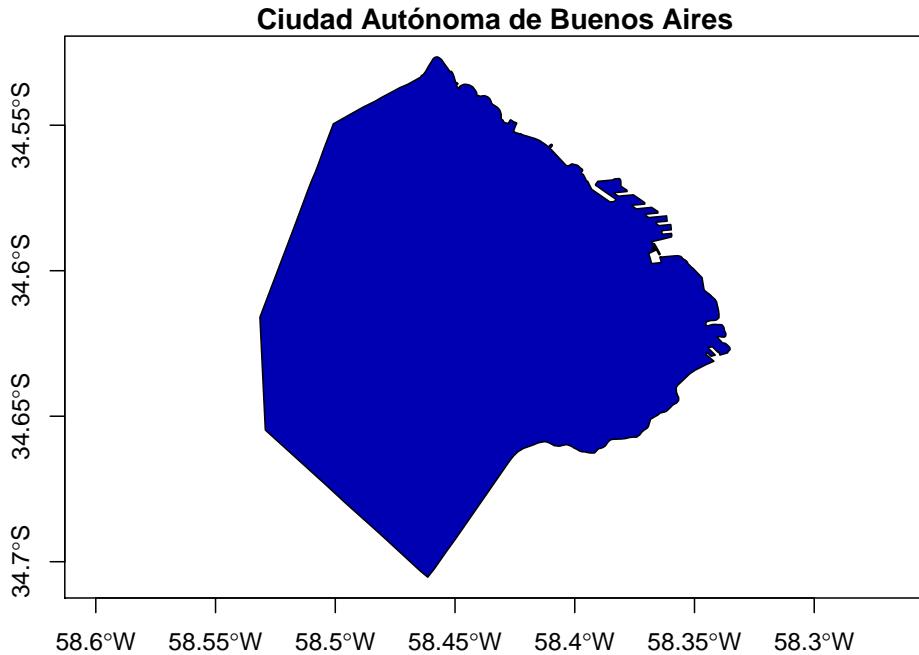
```
df$fna <- gsub(pattern = "Provincia ", replacement = "", x = df$fna)
```

La siguiente linea de codigo es muy barbara / da hora / cool / la raja. Corran!

```
library(mapview)  
mapview(df, zcol = "fna", legend = T, alpha = 0.8)
```

Ahora vamos a seleccionar Buenos Aires

```
plot(df[df$fna == unique(df$fna)[1], "entidad"],  
      axes = T,  
      main = unique(df$fna)[1])
```



Ahora, vamos a extraer los centroids de cada poligono

Recuerda, si tienes dudas de alguna funcion, clica ?FUNCION, por ejemplo, ?st_centeroid

```
cent <- st_centroid(df)
```

El warning se debe a que los datos que tenemos estan en latitud y longitud, y por lo visto, tenemos un dato con limites que no son correctos. Me parece que es la Antartida, debido se ve demasiado grande. Vamos a sacarla y ver como da

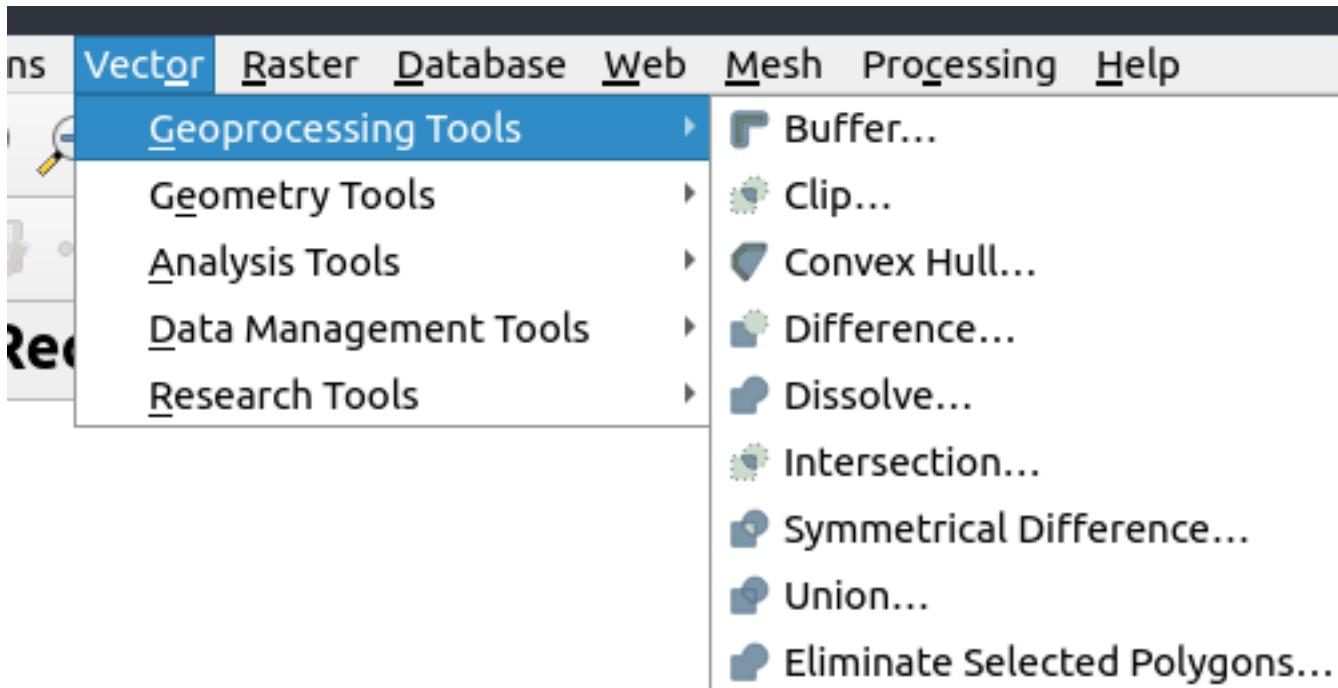
```
cent <- st_centroid(df[df$fna != "Antartida", ])
```

Ahora vemos que tenemos un warning diferente. En este caso, `st_centroid` nos avisa que los geograficos son constantes, pero como sabemos, la tierra es un elipsoide, no una esfera perfecta.

3.4 Geoprocесamiento

El geoprocесamiento incluye diferentes tareas como buffers y otros, como se ve en QGIS

```
knitr::include_graphics("figs/geo.png")
```



Iremos a calcular el buffer de cada centroid, sin embargo, primeros vamos a proyectar nuestros datos para UTM con un usado codigo llamado EPSG (European Petroleum Survey Group), como muestra este blog cafe geodesico. Asi cada provincia podria tener un EPSG diferente, para simplificar este curso, utilizaremos el codigo epsg 5346.

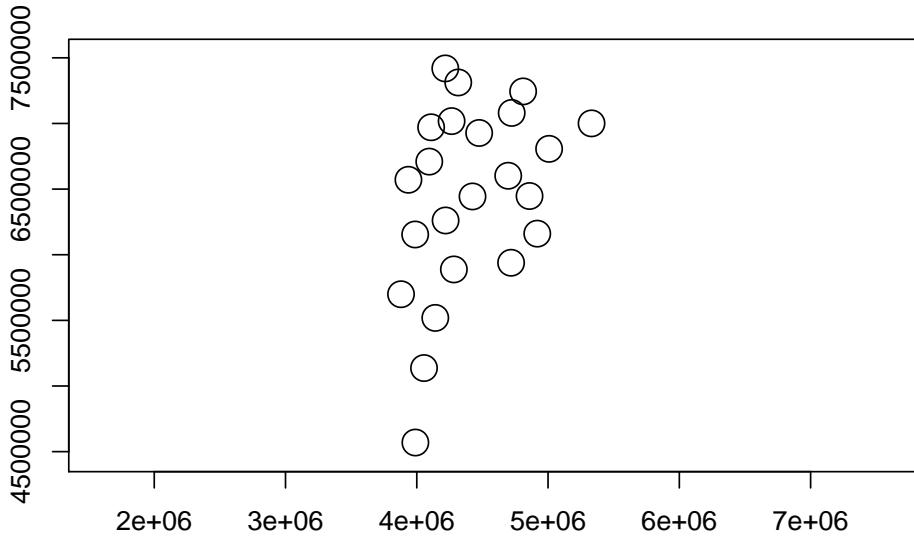
Para transformar nuestros datos, tenemos que:

```
dft <- st_transform(cent, 5346)
dft
```

```
## Simple feature collection with 23 features and 8 fields
## geometry type:  POINT
## dimension:      XY
## bbox:            xmin: 3879876 ymin: 4569475 xmax: 5331263 ymax: 7419288
## epsg (SRID):   5346
## proj4string:    +proj=tmerc +lat_0=-90 +lon_0=-63 +k=1 +x_0=4500000 +y_0=0 +ellps=GRS80
## # A tibble: 23 x 9
##       entidad objeto fna   gna   nam   in1   fdc   sag
##       <int> <chr> <chr> <chr> <chr> <chr> <chr>
## 1          0 Provi~ Ciud~ Ciud~ 02   Geog~ IGN
## 2          0 Provi~ del ~ Prov~ Neuq~ 58   Geog~ IGN
## 3          0 Provi~ de L~ Prov~ La P~ 42   Geog~ IGN
## 4          0 Provi~ de M~ Prov~ Mend~ 50   Geog~ IGN
## 5          0 Provi~ de S~ Prov~ San ~ 74   Geog~ IGN
## 6          0 Provi~ de C~ Prov~ Córd~ 14   Geog~ IGN
## 7          0 Provi~ de S~ Prov~ Sant~ 82   Geog~ IGN
## 8          0 Provi~ de E~ Prov~ Entr~ 30   Geog~ IGN
## 9          0 Provi~ de S~ Prov~ San ~ 70   Geog~ IGN
## 10         0 Provi~ de L~ Prov~ La R~ 46   Geog~ IGN
## # ... with 13 more rows, and 1 more variable: geometry <POINT [m]>
```

Luego, calculamos el buffer, que puede tener una distancia constante, o una distancia variable:

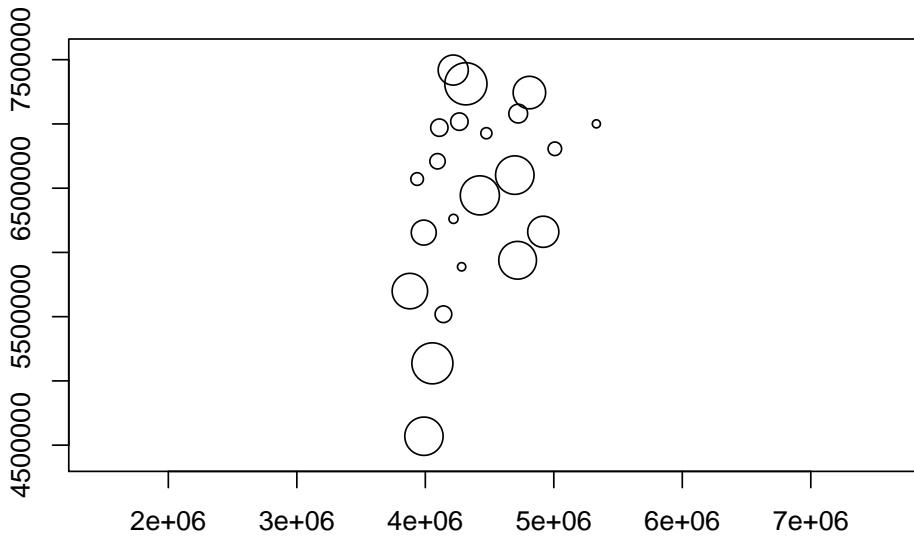
```
dft_pro <- st_buffer(x = dft, dist = 100000) # 100 km
plot(dft_pro$geometry, axes = T)
```



Macanudo /Bacan / barbaro / Chevere / Da Ora / Splendid / Cool!

Podemos calcular perfectamente un buffer con una distancia variable.....

```
dft_pro <- st_buffer(x = dft,
                      dist = rnorm(n = nrow(dft), mean = 100000, sd = 50000)) # 100 km
plot(dft_pro$geometry, axes = T)
```



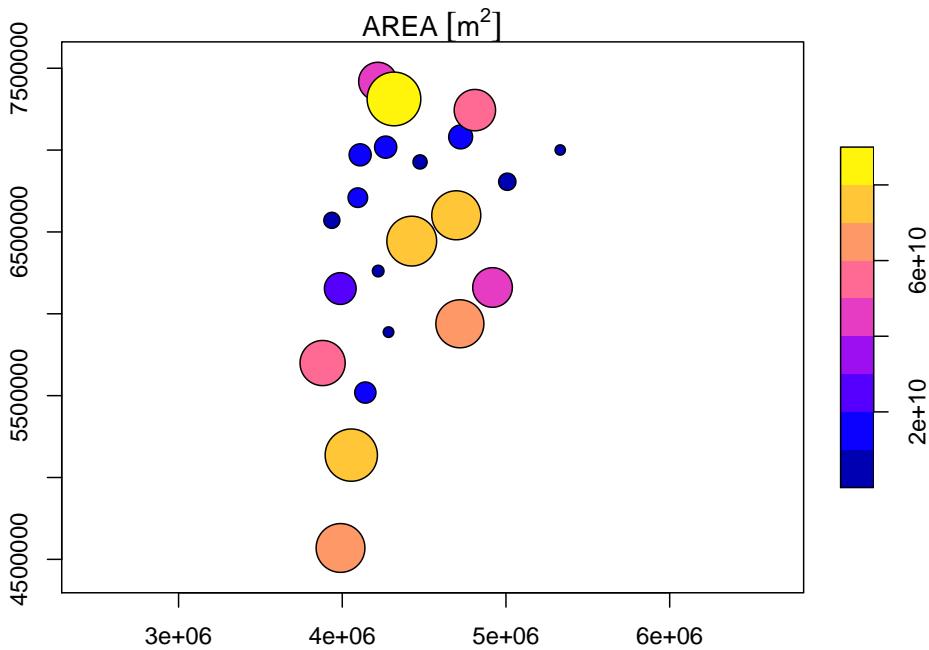
Con este grafico hemos ploteado un buffer que tiene un comprimento aleatorio pero con una media de 100 km y un desvio padrao de 50 km, para marcar bien las posibilidades....

Antes de terminar, vamos plotear el area de cada buffer. Como el diametro fue

56CHAPTER 3. TALLER VECTORES: APLICACIÓN DE SOFTWARE DE INFORMACIÓN GEOGRÁFICA

calculado de forma aleatoria, en cada rodada este numero cambia, entonces:

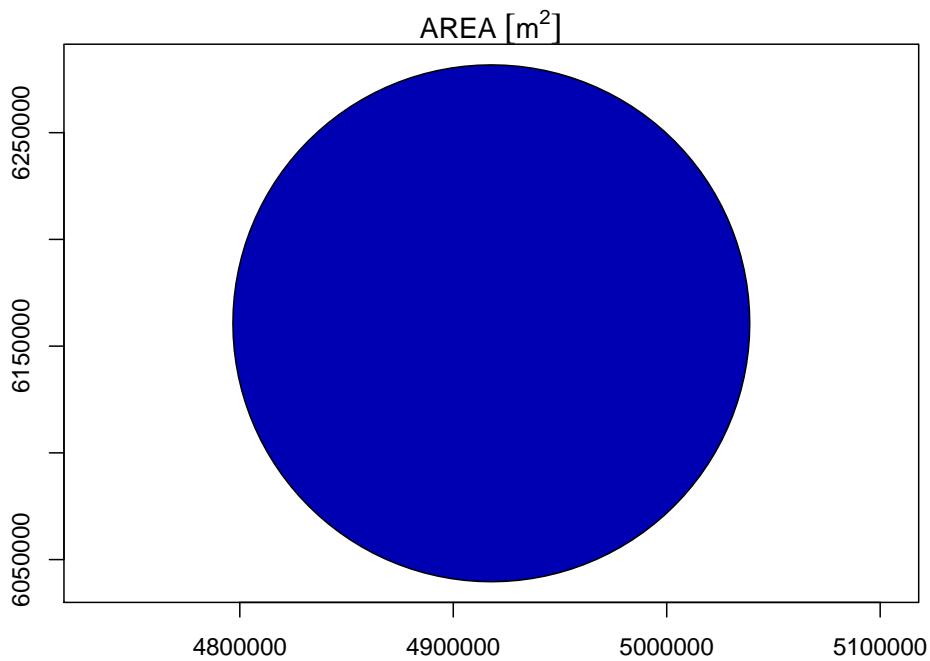
```
dft_pro$AREA <- st_area(dft_pro)
plot(dft_pro["AREA"], axes = T) # Area em m^2
```



Honestamente, puedes hacer muuuchas cosas.

Vamos a plotear la primera linea

```
dft_pro1 <- dft_pro[1, ] #primera linea
plot(dft_pro1["AREA"], axes = T) # Area em m^2
```



lo que mas me ha asombrado / shocked / chocado / con sf, es que el processo de gridding de VEIN en qgis demoraba 10 minutos, pero usando sf rn R, demora **5.5 segundos**

Pero ahora vamos a leer datos meteorologicos y climaticos en grillas (RASTER)
(next chapter)

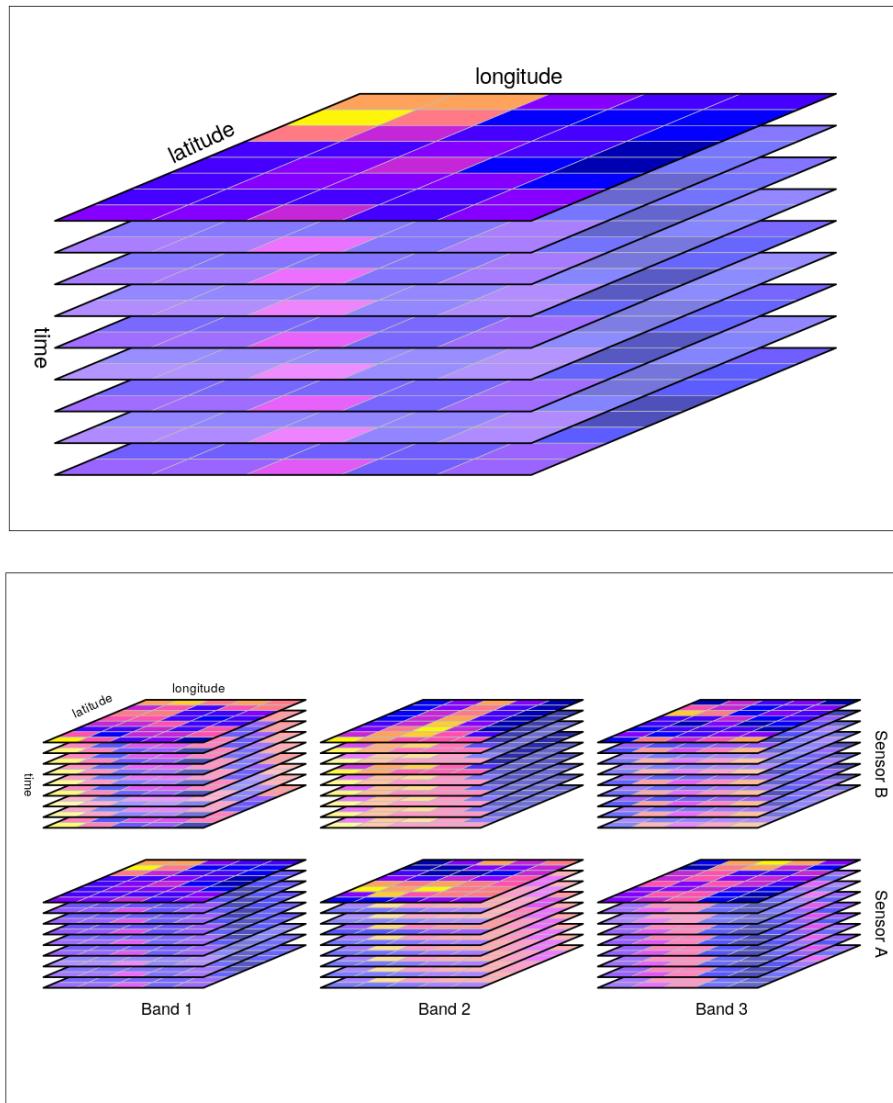
Chapter 4

Taller RASTER Y CUBOS DE DATOS VECTORIALES: Aplicación de software de información geográfica y modelado

gracias Amanda Rehbein <3 (amanda.rehbein@usp.br)

Raster son informacion espaciales en una grilla espacial. En general, la informacion de las frillas tiene una distancia homogenea, Por ejemplo, vea las siguientes figuras:

60CHAPTER 4. TALLER RASTER Y CUBOS DE DATOS VECTORIALES: APLICACIÓN DE SOFTWARE



Por ejemplo, cada banda podria representar un tiempo. Como se ve, estos no en realidad son cubos de datos de varias dimensiones, sin embargo, pueden existir grillas irregulares, como las que osn bien representadas con el paquete stars

Preparense, porque vamos a avanzar directamente en el fondo del mundo de los datos espacio/temporales grillados raster y stars

raster es una clase y una libreria de R que representa datos en grillas, llamados raster **stars** es una clase y una libreria de R que representa datos en grillas, llamados tambien rasters. Por lo tanto, **stars** y **raster** tienen conexiones para leer los datos en formato GDAL raster. Que quiere decir esto????

Quiere decir que, tal como GDAL lee vectores, tambien lee [raster]((<https://gdal.org/drivers/raster/index.html>)). Como por ejemplo:

- ASCII
- ENVI
- GeoTIFF
- JAXA PALSAR
- NetCDF
- GeoSpatial PDF
- R
- R Raster
- XYZ

Vamos meterle manos a la obra, vamos a nadar en el mundo multidimensional de los raster y si no saben nadar aprenderan como los pajaros que aprenden ante la subita e inherente necesidad vital de sobrevivir.

Antes de abrir los datos con los programas, vamos a ver que vamos tenemos

```
library(ncdf4)
nc_open("dados/pr_Amon_HadGEM2-ES_historical_r1i1p1_DJF_1985-2004.nc")

## File dados/pr_Amon_HadGEM2-ES_historical_r1i1p1_DJF_1985-2004.nc (NC_FORMAT_CLASSIC):
##
##      4 variables (excluding dimension variables):
##          double lon_bnds[bnnds,lon]
##          double lat_bnds[bnnds,lat]
##          double time_bnds[bnnds,time]
##          double pr[lon,lat,time]
##              standard_name: precipitation_flux
##              long_name: Precipitation
##              units: kg m-2 s-1
##              _FillValue: 1.00000002004088e+20
##              missing_value: 1.00000002004088e+20
##              comment: at surface; includes both liquid and solid phases from all types of clouds
##              original_name: mo: m01s05i216
##              history: 2011-03-22T15:05:23 out-of-bounds adjustments: (-1e-07 <= x < 0.0) => set
##              cell_methods: time: mean
##              associated_files: baseURL: http://cmip-pcmdi.llnl.gov/CMIP5/dataLocation gridspecF
##
##      4 dimensions:
##          lon  Size:192
##              standard_name: longitude
##              long_name: longitude
##              units: degrees_east
##              axis: X
##              bounds: lon_bnds
##          bnnds  Size:2
```

62CHAPTER 4. TALLER RASTER Y CUBOS DE DATOS VECTORIALES: APLICACIÓN DE SOFTWARE

```
##      lat  Size:145
##            standard_name: latitude
##            long_name: latitude
##            units: degrees_north
##            axis: Y
##            bounds: lat_bnds
##      time  Size:19    *** is unlimited ***
##            standard_name: time
##            long_name: time
##            bounds: time_bnds
##            units: days since 1859-12-1 00:00:00
##            calendar: 360_day
##            axis: T
##
##      30 global attributes:
##      CDI: Climate Data Interface version ?? (http://mpimet.mpg.de/cdi)
##      Conventions: CF-1.4
##      history: Sun Sep 08 00:37:13 2019: cdo selmon,1 tmp_SELTIME_SEASSUM_pr_Amon
##      Sun Sep 08 00:36:18 2019: cdo selyear,1986,1987,1988,1989,1990,1991,1992,1993,1994,
##      Sun Sep 08 00:33:38 2019: cdo -b 64 seassum pr_Amon_HadGEM2-ES_historical_r1i1p1_19
##      MOHC pp to CMOR/NetCDF convertor (version 1.7.0) 2011-03-22T15:05:17Z CMOR rewrote
##      source: HadGEM2-ES (2009) atmosphere: HadGAM2 (N96L38); ocean: HadGOM2 (lat
##      institution: Met Office Hadley Centre, Fitzroy Road, Exeter, Devon, EX1 3PB
##      institute_id: MOHC
##      experiment_id: historical
##      model_id: HadGEM2-ES
##      forcing: GHG, SA, Oz, LU, S1, V1, BC, OC, (GHG = CO2, N2O, CH4, CFCs)
##      parent_experiment_id: piControl
##      parent_experiment_rip: r1i1p1
##      branch_time: 0
##      contact: chris.d.jones@metoffice.gov.uk, michael.sanderson@metoffice.gov.uk
##      references: Bellouin N. et al, (2007) Improved representation of aerosols f
##      initialization_method: 1
##      physics_version: 1
##      tracking_id: 15cb947c-34d2-4da2-b830-9dd0123d96d3
##      mo_runid: ajhoh
##      product: output
##      experiment: historical
##      frequency: mon
##      creation_date: 2011-03-22T19:30:17Z
##      project_id: CMIP5
##      table_id: Table Amon (06 January 2011) 085e3f5c098d203b2fd5d296226f41b3
##      title: HadGEM2-ES model output prepared for CMIP5 historical
##      parent_experiment: pre-industrial control
##      modeling_realm: atmos
##      realization: 1
```

```
##      cmor_version: 2.5.3
##      CDO: Climate Data Operators version 1.7.2 (http://mpimet.mpg.de/cdo)
```

ERA5!

```
library(stars)
pcp_era5 <- read_stars("dados/pr_Amon_HadGEM2-ES_historical_r1i1p1_DJF_1985-2004.nc")
names(pcp_era5)
str(pcp_era5)
head(pcp_era5)
```

brick es como literalmente un ladrillo considerndo usualmente la dimension temporal como un ladrillo.

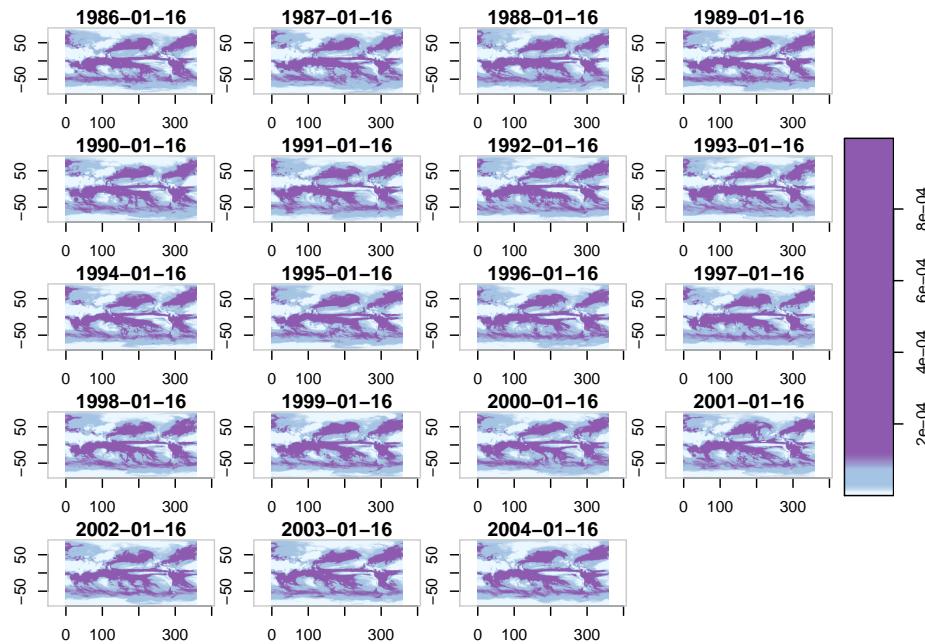
como vemos, la funcion `str` entrega mucha informacion. Esta funcion no es comunmente usada, pero su archivo de documentacion es `?str`

ERA5, investiguen!

Vamos a plotear estos datos!

```
library(stars)
library(cptcity)
pcp_era5 <- read_stars("dados/pr_Amon_HadGEM2-ES_historical_r1i1p1_DJF_1985-2004.nc")
plot(pcp_era5,
     col = lucky(),
     axes = TRUE)

## Colour gradient: cb_seq_BuPu_03, number: 453
```



Ahora vamos a mostrar todo un script para hacer graficos y analizar datos climáticos usando la librería **raster**

```
# -----#
# Precipitacion y temperatura medias #
# Verano y invierno #
# (very simple analysis) #
# -----#

library(raster)
library(cptcity)

path <- "dados"

# DJF - VERANO: ####

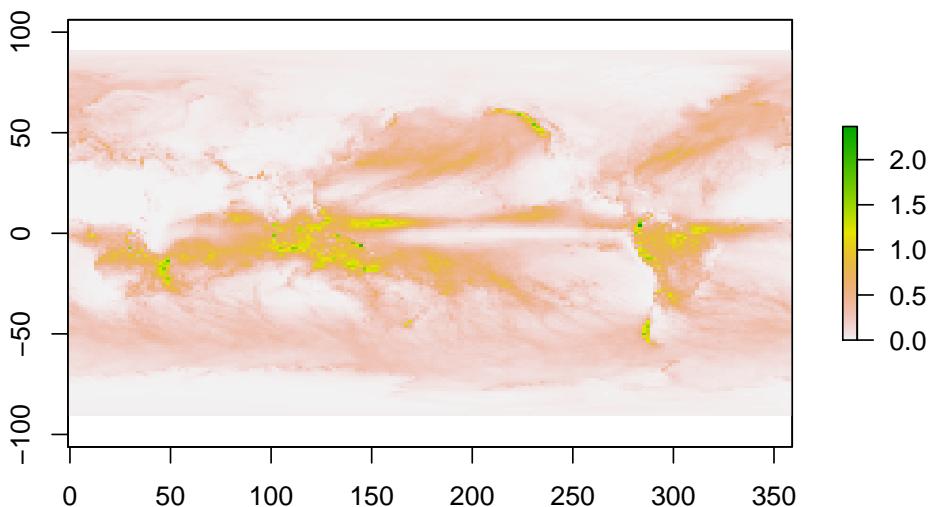
# Open the precipitation file:
pcp_era5 <- brick("dados/r_ERA5_pcp_DJF_1985-2004.nc")
# Print information and metadata:
pcp_era5 # This file has 17 layers (years)

## class      : RasterBrick
## dimensions : 145, 192, 27840, 21 (nrow, ncol, ncell, nlayers)
## resolution : 1.875, 1.25 (x, y)
## extent     : -0.9375, 359.0625, -90.625, 90.625 (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
```

```

## source      : /home/sergio/Documents/UBA/dados/r_ERA5_pcp_DJF_1985-2004.nc
## names      : X1984.01.16.12.06.28, X1985.01.01.00.06.28, X1986.01.01.01.06.28, X1987.01.01.01.
## Date/time   : 1984-01-16 12:06:28, 2005-12-01 01:06:28 (min, max)
## varname     : mtp
# Plot the 1st time/year:
plot(pcp_era5[[1]]*3600)      # *3600 to obtain mm/h

```



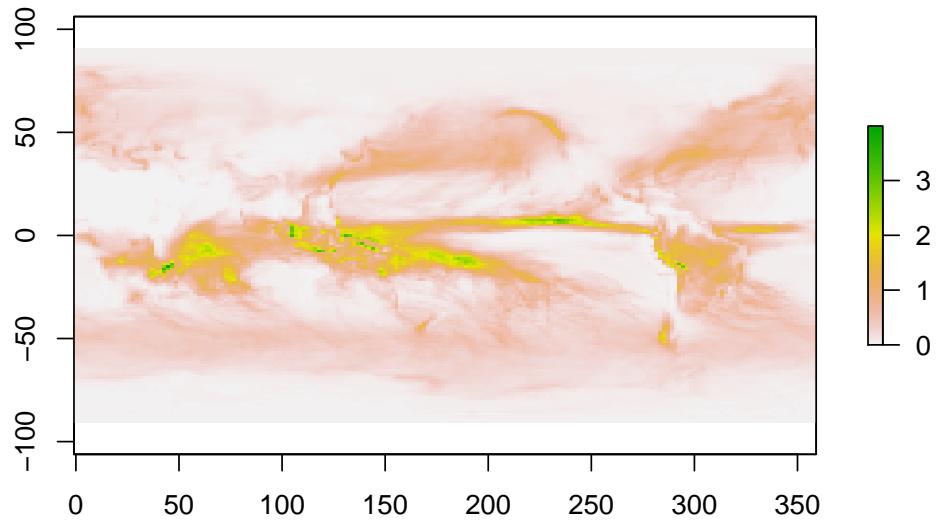
```

# Calculate the mean:
mean_pcp_era <- mean(pcp_era5*3600)

# Opening CMIP5 data model: HadGEM2-ES\
pcp_historical <- brick("dados/pr_Amon_HadGEM2-ES_historical_r1i1p1_DJF_1985-2004.nc")
plot(pcp_historical[[1]]*3600)

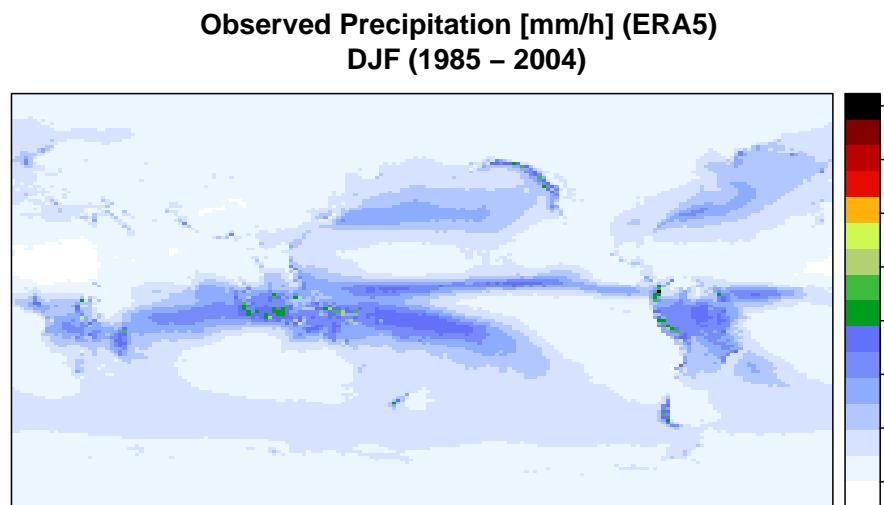
```

66 CHAPTER 4. TALLER RASTER Y CUBOS DE DATOS VECTORIALES: APLICACIÓN DE SOFTWARE



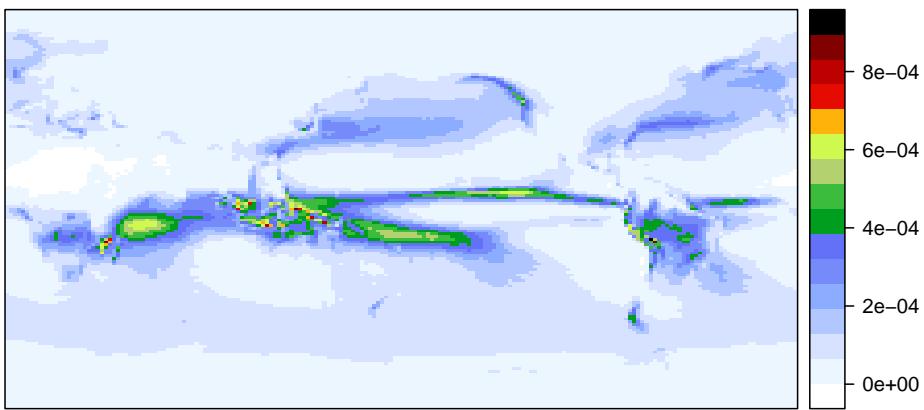
```
# calculando a media
mean_pcp_historical <- mean(pcp_historical)

spplot(mean_pcp_era, interpolate = T,
       main = "Observed Precipitation [mm/h] (ERA5) \n DJF (1985 - 2004)",
       col.regions = cpt("ncl_precip2_17lev"))
```



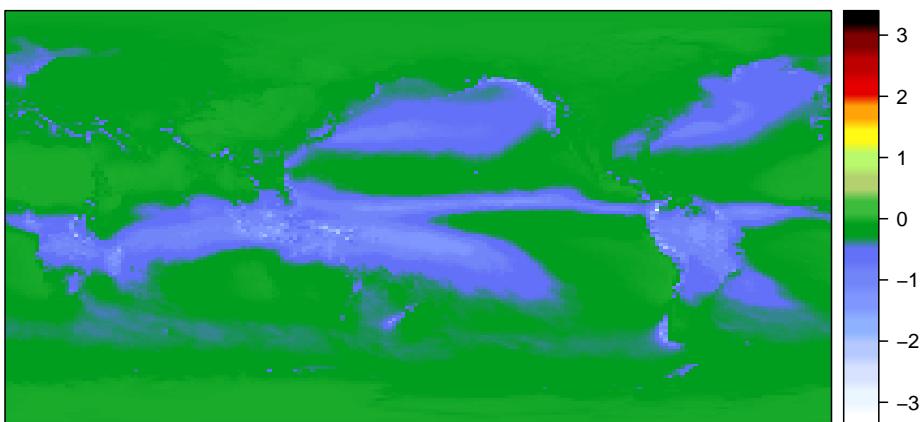
```
spplot(mean_pcp_historical,
       main = "Simulated precipitation [mm/h] (HadGEM2-ES) \n DJF (1985 - 2004)",
       col.regions = cpt("ncl_precip2_17lev"))
```

Simulated precipitation [mm/h] (HadGEM2-ES)
DJF (1985 – 2004)



```
spplot(mean_pcp_historical - mean_pcp_era,
       main = "Simulated - Observed precipitation [mm/h] \n DJF (1985 - 2004)",
       col.regions = cpt("ncl_precip2_17lev", n = length(seq(-3.4, 3.4, 0.01))),
       at= seq(-3.4, 3.4, 0.01))
```

Simulated – Observed precipitation [mm/h]
DJF (1985 – 2004)



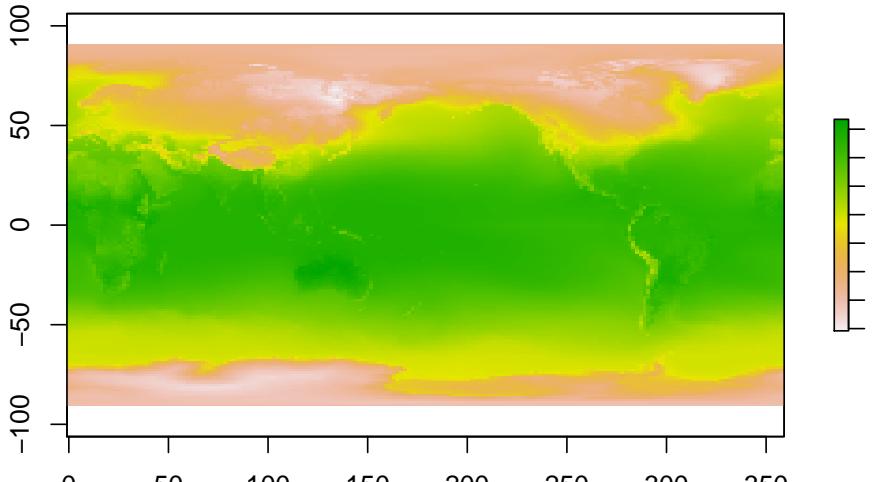
Ahora vamos a ver la temperatura

```
# Temperature ####
t2m_era5 <- brick("dados/r ERA5_t2m_DJF_1985-2004.nc")
t2m_era5

## class      : RasterBrick
## dimensions : 145, 192, 27840, 17  (nrow, ncol, ncell, nlayers)
## resolution : 1.875, 1.25  (x, y)
```

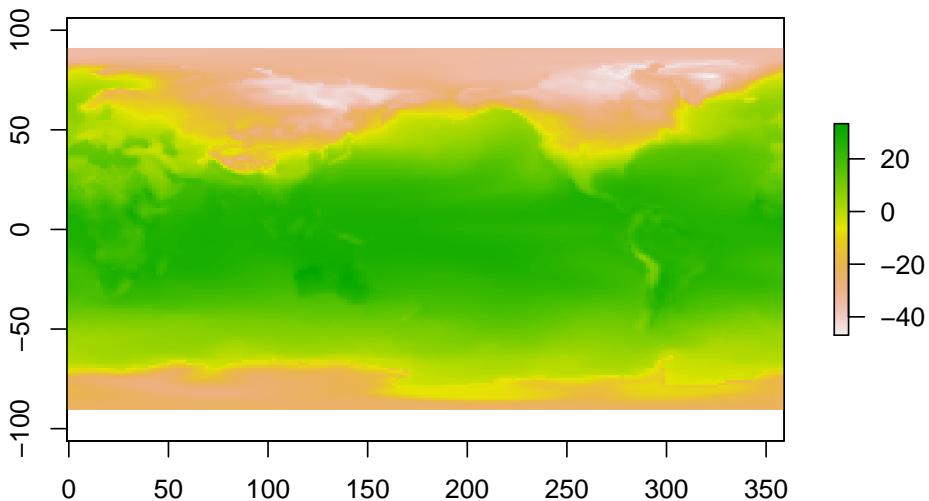
68CHAPTER 4. TALLER RASTER Y CUBOS DE DATOS VECTORIALES: APLICACIÓN DE SOFTWARE

```
## extent      : -0.9375, 359.0625, -90.625, 90.625 (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## source     : /home/sergio/Documents/UBA/dados/r_ERA5_t2m_DJF_1985-2004.nc
## names      : X1986.01.01.01.06.28, X1987.01.01.01.06.28, X1988.01.01.01.06.28, X1989.01.01.01.06.28
## Date/time   : 1986-01-01 01:06:28, 2004-01-01 01:06:28 (min, max)
## varname    : t2m
plot(t2m_era5[[1]] - 273.15)      # -273.15 to obtain °C
```



```
mean_t2m_era <- mean(t2m_era5 - 273.15)
```

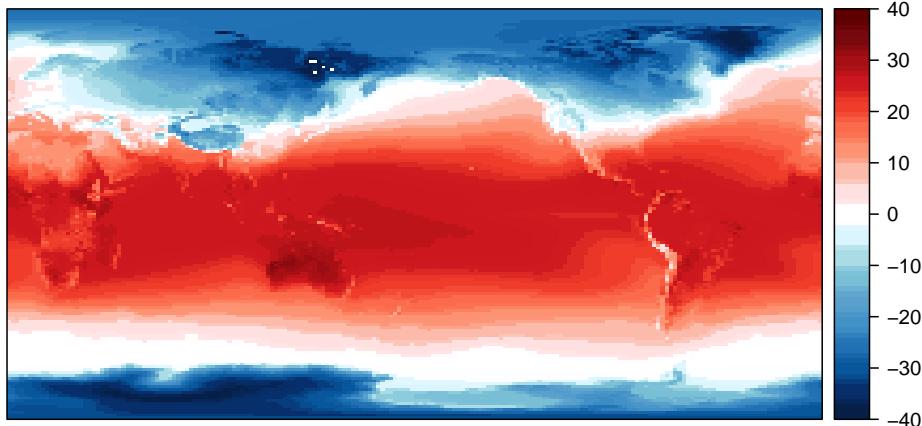
```
t2m_historical <- brick("dados/tas_Amon_HadGEM2-ES_historical_r1i1p1_DJF_1985-2004.nc")
plot(t2m_historical[[1]] - 273.15)
```



```
mean_t2m_historical <- mean(t2m_historical - 273.15)

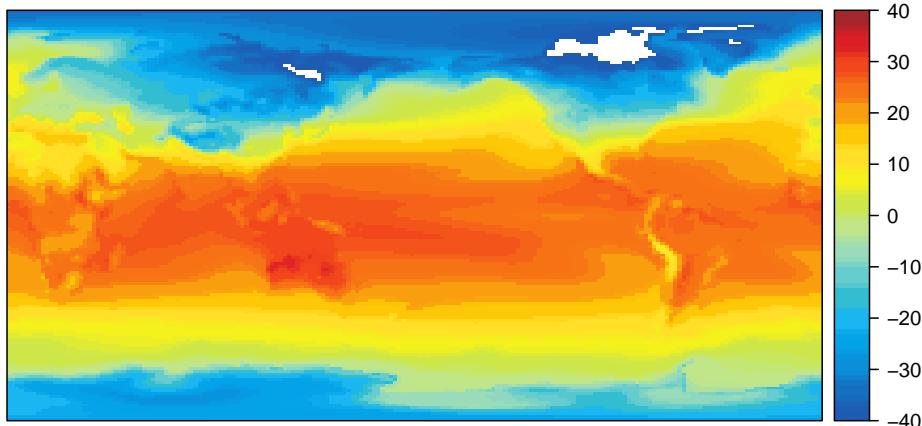
spplot(mean_t2m_era,
       main = "Observed temperature at 2m [°C] (ERA5) \n DJF (1985 - 2004)",
       at = seq(-40, 40, 1),
       col.regions = cpt("ncl_temp_diff_18lev"))
```

**Observed temperature at 2m [°C] (ERA5)
DJF (1985 – 2004)**



```
spplot(mean_t2m_historical,
       main = "Simulated temperature at 2m [°C] (HadGEM2-ES) \n DJF (1985 - 2004)",
       at = seq(-40, 40, 1),
       col.regions = cpt("arendal_temperature"))
```

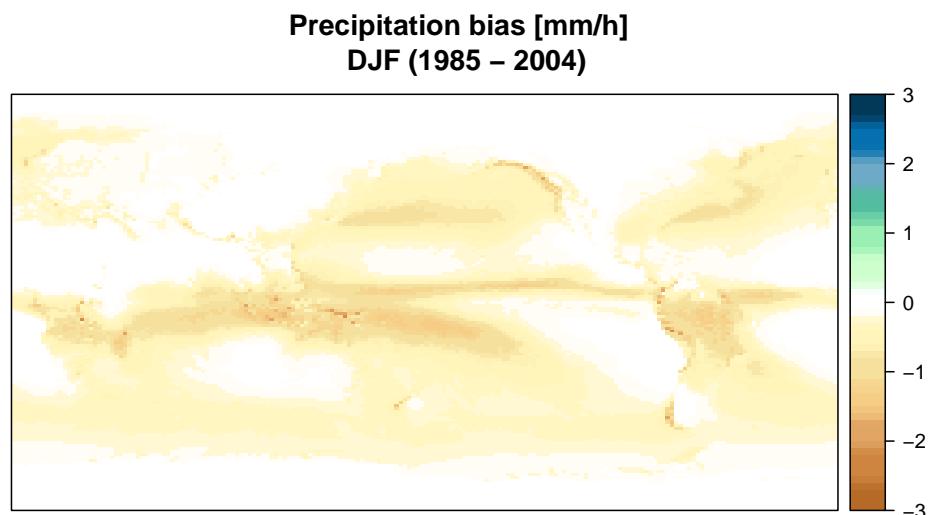
**Simulated temperature at 2m [°C] (HadGEM2-ES)
DJF (1985 – 2004)**



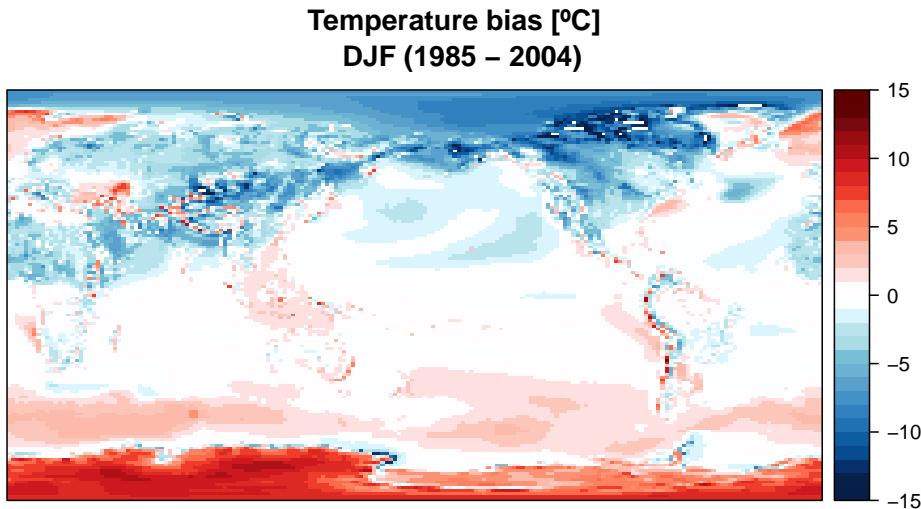
70CHAPTER 4. TALLER RASTER Y CUBOS DE DATOS VECTORIALES: APLICACIÓN DE SOFTWARE

```
# Calculate the bias (how model is different from the observation):
pcp_bias <- (mean_pcp_historical - mean_pcp_era)
t2m_bias <- (mean_t2m_historical - mean_t2m_era)

spplot(pcp_bias,
       main = "Precipitation bias [mm/h] \n DJF (1985 - 2004)",
       at = seq(-3, 3, 0.1),
       col.regions = cpt("ncl_precip_diff_12lev"))
```



```
spplot(t2m_bias,
       main = "Temperature bias [°C] \n DJF (1985 - 2004)",
       at = seq(-15, 15, 1),
       col.regions = cpt("ncl_temp_diff_18lev"))
```



Y ahora vamos a ver como plotear datos de satelite de Sentinel 5P

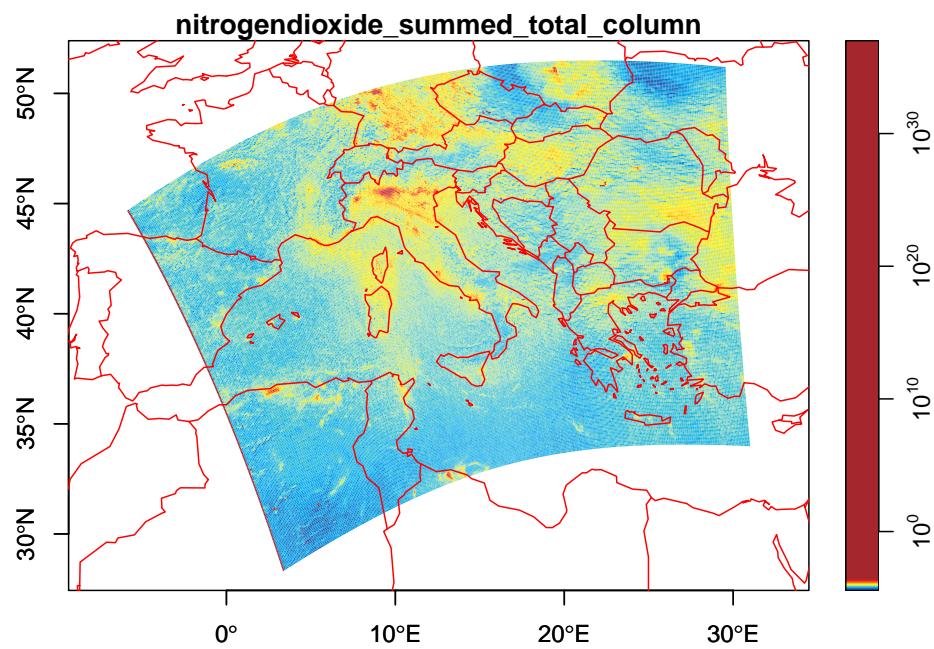
```
library(stars)
(s5p = system.file("sentinel5p/S5P_NRTI_L2__N02____20180717T120113_20180717T120613_03932_01_01000

## [1] "/home/sergio/R/x86_64-pc-linux-gnu-library/3.6/starsdata/sentinel5p/S5P_NRTI_L2__N02____2
nit.c = read_stars(s5p, sub = "//PRODUCT/SUPPORT_DATA/DETAILED_RESULTS/nitrogendioxide_summed_tot
                  curvilinear = c("//PRODUCT/longitude",
                                "//PRODUCT/latitude"),
                  driver = NULL)

## //PRODUCT/longitude,
## //PRODUCT/latitude,
## //PRODUCT/SUPPORT_DATA/DETAILED_RESULTS/nitrogendioxide_summed_total_column,
```

Y el plot es

```
plot(nit.c,
      breaks = "hclust",
      reset = FALSE,
      axes = TRUE,
      logz = TRUE,
      key.length = 1,
      col = cpt("arendal_temperature"))
maps::map('world', add = TRUE, col = 'red')
```



Chapter 5

Estudios de series de tiempo para asociar salud y variables ambientales

Uno de los metodos mas validados y establecidos para identificar asociaciones entre mortalidad, morbilidad, y otros parametros con variables ambientales son las series de tiempo.

Algunos de los autores clasicos sobre este tipo de estudios estan:

- Francesca Dominici [[link google scholar](#)]
- Luis Cifuentes [[link google scholar](#)]
- Petros Koutrakis [[link google scholar](#)]
- Paulo Saldíva [[link google scholar](#)]
- Michelle Bell [[link google scholar](#)]
- Antonela Zanometti [[link google scholar](#)]

Francesca Dominici ha publicado diferentes estudios y actualizado la metodologia de estudios de series de tiempo considerando la distribucion estadistica, control de los factores de confusión y modelos generales aditivos.

5.1 Distribucion

Los estudios de series de tiempo consideran en general variables diarias. En este sentido, las variables de salud son conteos diarios con una distribucion apropiada, generalmente Poisson o Binomial Negativa, por ejemplo. Estas son diferentes de la distribucion normal.

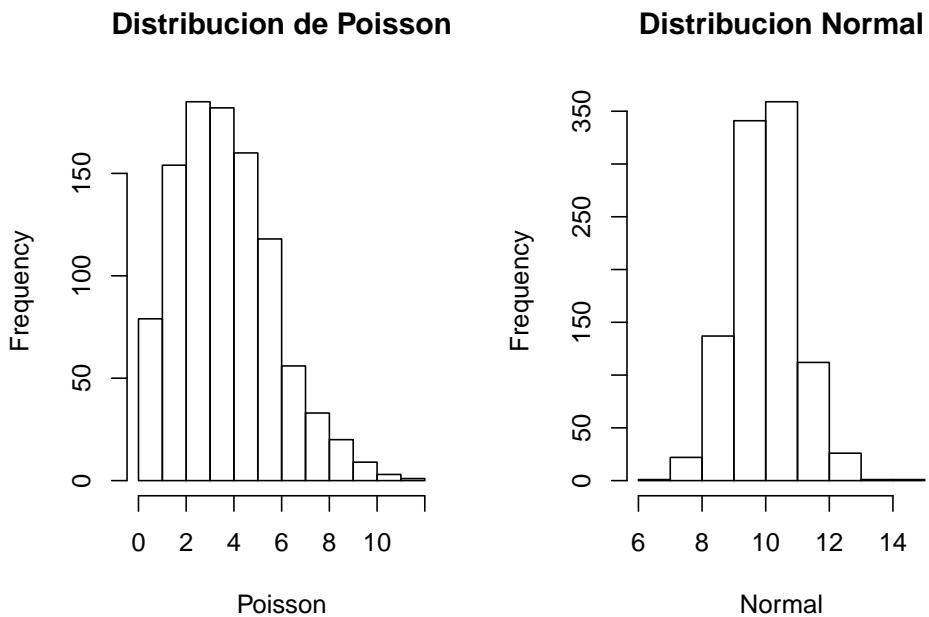
Densidad de la distribucion de Poisson:

$$p(x) = \lambda^x \exp(-\lambda)/x! \quad (5.1)$$

Densidad de la distribucion de Normal:

$$f(x) = 1/((2\pi)\sigma) e^{-((x-\mu)^2/(2\sigma^2))} \cdot 0.5 \quad (5.2)$$

Vamos a ver un ejemplo:



5.2 Efecto de la contaminacion

Dominici presento el modelo semi-parametrico usando modelos generales aditivos (Wood, 2017) para el estudio de series de tiempo. Este modelo consiste en:

$$\log(m) = \beta_0 + \beta_1 \cdot X_i + s(meteo) + s(time) + s(season) + \dots \quad (5.3)$$

Donde m es la variable dependiente sobre salud, β son los coeficientes, X_i son los contaminantes, s es una funcion spline

Luego, la asociacion resultante es el riesgo relativo (Abrutzky et al., 2013):

$$RR = \exp(\beta_i \cdot \delta X_i) \quad (5.4)$$

Que indica el riesgo relativo dada un cambio en la variable, por ejemplo, si se incrementan las concentraciones de PM10 en 50 ug/m³, el riesgo de mortalidad respiratoria aumenta 7%. Aplicando este porcentaje sobre la poblacion es posible calcular la mortalidad anual asociada a la contaminacion. Vamos a ver valores reales para Argentina y Chile (Abrutzky et al., 2013).

5.3 Efecto de la temperatura

Puede se utilizada la misma metodologia pero considerando variables meteorologicas en este caso. Un trabajo muy interesante recientemente publicado por el grupo de Paulo Saliva es el efecto de la temperatura en diferentes ciudades del mundo (Gasparrini et al., 2015).

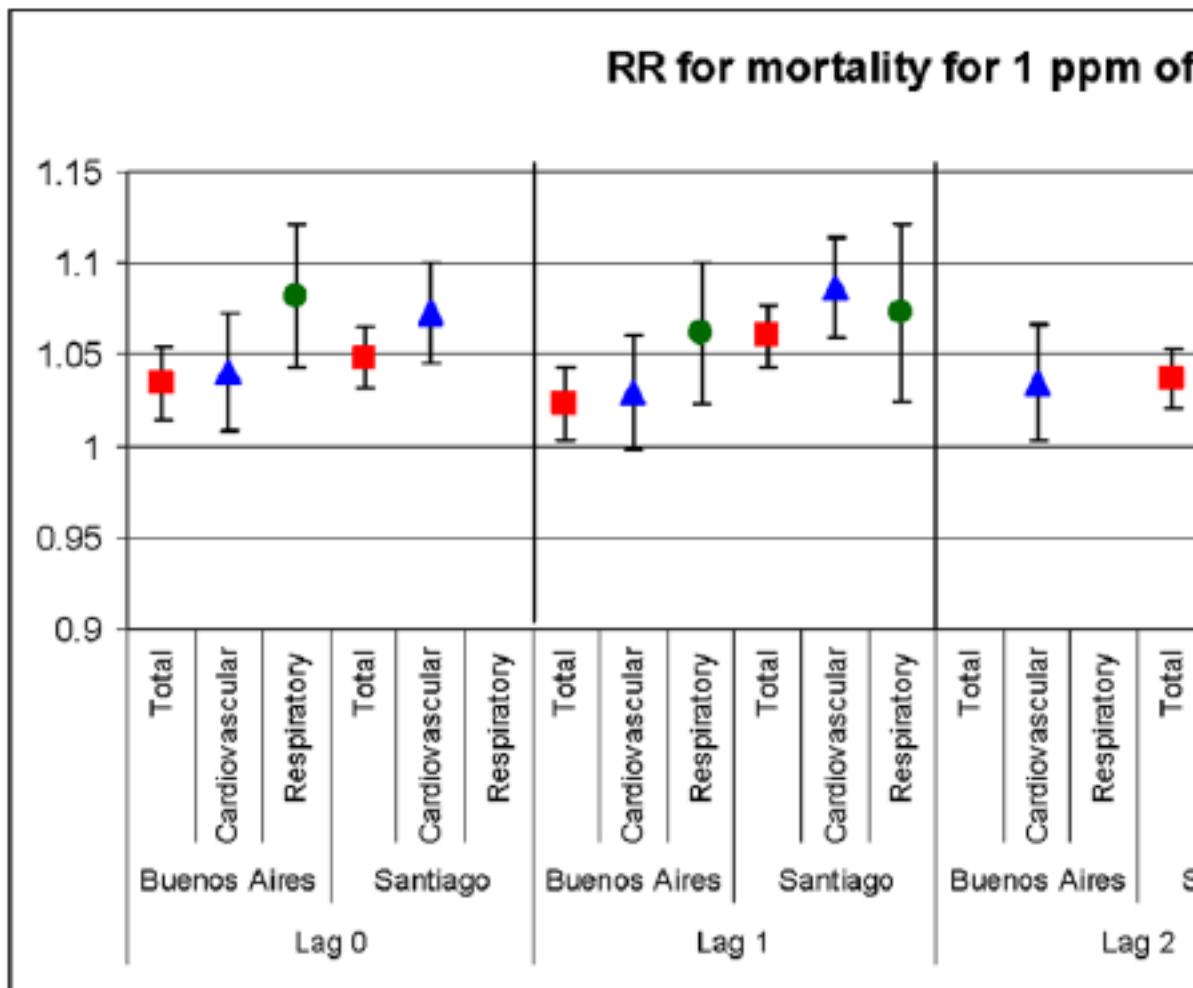
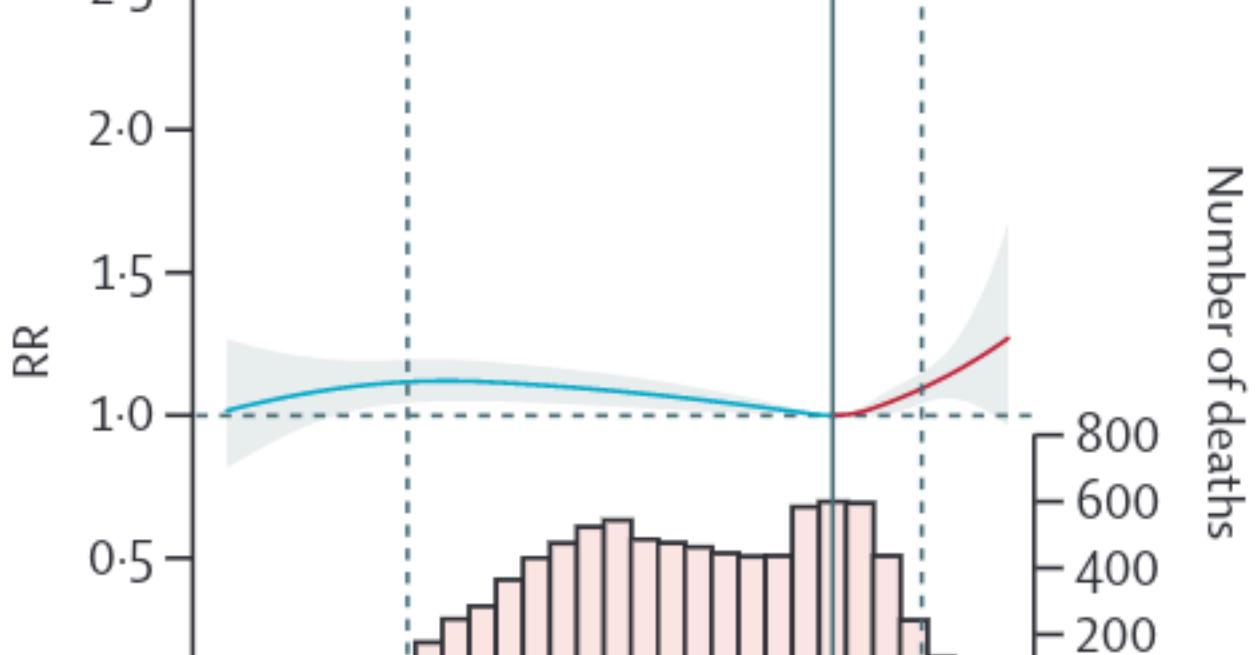
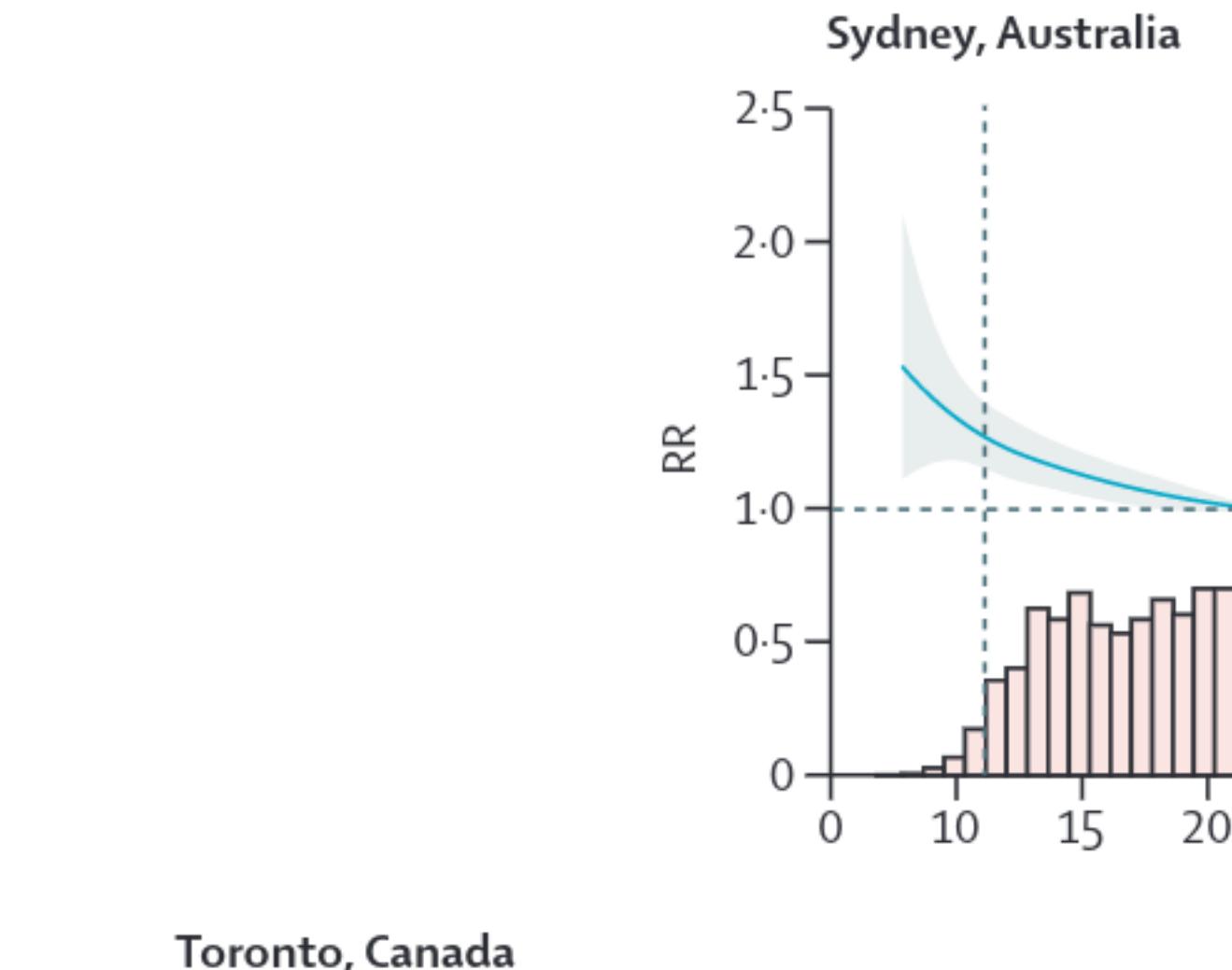


Figure 5.1: Riesgo relativo en mortalidad por incremento de 1 ppm de CO
(Abrutzky et al., 2013)



Bibliography

- Abrutzky, R., Ibarra-Espinosa, S., Matus, P., Romero-Lankao, P., Pereyra, V., and Dawidowski, L. (2013). Atmospheric pollution and mortality. a comparative study between two latin american cities: Buenos aires (argentina) and santiago (chile). *Int. J. Environment and Health*, 6.
- Bell, M. L. and Davis, D. L. (2001). Reassessment of the lethal london fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environmental health perspectives*, 109(suppl 3):389–394.
- Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., et al. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552.
- Brasseur, G., Orlando, J. J., Tyndall, G. S., et al. (1999). *Atmospheric chemistry and global change*. Oxford University Press.
- Brasseur, G. P. and Jacob, D. J. (2017). *Modeling of atmospheric chemistry*. Cambridge University Press.
- Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R., Verheggen, B., Maibach, E. W., Carlton, J. S., Lewandowsky, S., Skuce, A. G., Green, S. A., et al. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, 11(4):048002.
- Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., Leone, M., Sario, M. D., Bell, M. L., Guo, Y.-L. L., fu Wu, C., Kan, H., Yi, S.-M., de Sousa Zanotti Staglorio Coelho, M., Saldiva, P. H. N., Honda, Y., Kim, H., and Armstrong, B. (2015). Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet*, 386(9991):369 – 375.
- Hijmans, R. J. (2019). *raster: Geographic Data Analysis and Modeling*. R package version 3.0-2.
- Ibarra-Espinosa, S., Ynoue, R., O’Sullivan, S., Pebesma, E., Andrade, M. D. F., and Osses, M. (2018). Vein v0.2.2: an r package for bottom-up vehicular emissions inventories. *Geoscientific Model Development*, 11(6):2209–2229.

- Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., González Abad, G., Zhang, L., Zhang, X., Fried, A., Crounse, J. D., St. Clair, J. M., and Wisthaler, A. (2018). High-resolution inversion of omi formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast us. *Atmospheric Chemistry and Physics*, 18(8):5483–5497.
- Landrigan, P. J., Fuller, R., Acosta, N. J., Adeyi, O., Arnold, R., Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., et al. (2018). The lancet commission on pollution and health. *The Lancet*, 391(10119):462–512.
- Macedonio, G., Pareschi, M. T., and Santacroce, R. (1988). A numerical simulation of the plinian fall phase of 79 ad eruption of vesuvius. *Journal of Geophysical Research: Solid Earth*, 93(B12):14817–14827.
- Ntziachristos, L. and Samaras, Z. (2016). Emeep/eea emission inventory guide-book; road transport: Passenger cars, light commercial trucks, heavy-duty vehicles including buses and motorcycles. *European Environment Agency, Copenhagen*.
- OMS (2019). *Global Platform on Air Quality and Health*. <http://goo.gl/y0JhjI>.
- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446.
- Pebesma, E. (2019). *stars: Spatiotemporal Arrays, Raster and Vector Data Cubes*. <https://r-spatial.github.io/stars/>, <https://github.com/r-spatial/stars/>.
- Pulles, T. and Helsinga, D. (2010). *The art of emissions inventorying*. TNO.
- Schimel, D., Alves, D., Enting, I., Heimann, M., Joos, F., Raynaud, D., Wigley, T., Prather, M., Derwent, R., Ehhalt, D., et al. (1996). Radiative forcing of climate change. *Climate change 1995: The science of climate change*, pages 65–131.
- Stocker, T. (2014). *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Wikipedia contributors (2019a). Aerosol — Wikipedia, the free encyclopedia. <https://en.wikipedia.org/w/index.php?title=Aerosol&oldid=914421453>. [Online; accessed 9-September-2019].
- Wikipedia contributors (2019b). Great smog of london — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Great_Smog_of_London&oldid=914236507. [Online; accessed 8-September-2019].
- Wikipedia contributors (2019c). Great smog of london — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Great_Smog_of_London&oldid=914236507. [Online; accessed 8-September-2019].

- Wikipedia contributors (2019d). Mount vesuvius — Wikipedia, the free encyclopedia. [Online; accessed 8-September-2019].
- Wikipedia contributors (2019e). Nelson's column — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Nelson%27s_Column&oldid=914602869. [Online; accessed 8-September-2019].
- Wikipedia contributors (2019f). Ozone — Wikipedia, the free encyclopedia. <https://en.wikipedia.org/w/index.php?title=Ozone&oldid=914124258>. [Online; accessed 8-September-2019].
- Wikipedia contributors (2019g). Tropospheric ozone — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Tropospheric_ozone&oldid=911584695. [Online; accessed 8-September-2019].
- Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edition.