

TFM

# Cultivos y climatología. Situación Mundial.

*Sara Belén Ramos González*

Máster en Ciencia de Datos – UOC

Área 5 - Aula 1

Tutor: Rafael Luque Ocaña

PRA: Albert Solé



# Descripción y Justificación

## TFM Investigación



Afecciones en ***todos los ámbitos de nuestra sociedad***



Agricultura y clima



EFFECTOS  
NEGATIVOS  
DEVASTADORES



**RECUPERACIÓN**

*Personalmente*



Cuerpo  
Mente  
Naturaleza



**BIOTECNOLOGÍA**



# Estado del Arte

*Búsqueda  
pormenorizada de  
información*

- Extensa bibliografía
- Conclusiones contrastadas

- Fuentes de datos
- Variables
- Técnicas

- *Emisiones antropogénicas de gases de efecto invernadero (GEI) → (metano y óxido nitroso)*
- *La agricultura es culpable de la mayor parte de las emisiones.*
- *Afecciones en los patrones climáticos: temperaturas y precipitaciones.*



# Fuentes de datos

- **Contrastadas**
- **Fiables**
- **Robustas**

- Berkeley Earth
- Grupo Banco Mundial (The World Bank)
- Climate Change Knowledge Portal
- Eurostat
- Dryad
- Organización para la Cooperación y el Desarrollo Económico (OECD)

BERKELEY EARTH™





# La CCEG y los ODS



- Escasez de recursos
- Riesgos de contaminación
- Vulnerabilidad
- Condiciones de vida
- Modelo de producción y consumo
- Uso de tecnologías limpias y eficientes



# Datos y Variables en estudio

Años datos existentes:  
[1990 – 2023]



Años datos de predicción:  
[2024-250]

- ODS
- Pacto verde europeo



Variables:

- **Emisiones de gases** (mil toneladas métricas de equivalente de CO<sub>2</sub>): Óxido Nitroso (N<sub>2</sub>O) y Metano (CH<sub>4</sub>).
- **Uso de fertilizantes** (miles de toneladas de nutrientes): potásicos (K<sub>2</sub>O), nitrogenados (N) y fosfatados (P<sub>2</sub>O<sub>5</sub>).
- **Anomalía de Temperatura de la Tierra** (°C).
- **Población** (cantidad de personas).
- **Extensión de tierra cultivada** (sq. Km).
- **Precipitaciones** (mm).



# Preprocesamiento de datos

- ***Exploración datos obtenidos***
  - ✓ Descarga información
  - ✓ Entendimiento de la información
- ***Data wrangling y data cleansing***
  - ✓ Conversión a data frames
  - ✓ Conversión tipo de datos y formatos
  - ✓ Estandarización y homogeneización de nombres de variables.
  - ✓ Missing data → imputación/eliminación
  - ✓ Outliers
  - ✓ Union / Match
- ***Exportación de datos***



# Análisis de datos y visualización

- *Cálculos estadísticos*
- *Representaciones gráficas*
- *Aspectos estudiados sobre las variables:*
  - ✓ Descripción
  - ✓ Distribución
  - ✓ Correlación
  - ✓ Tendencias de comportamiento
- *Informe de visualización*
  - ✓ Dinámico
  - ✓ Orden y disposición
  - ✓ Tamaños
  - ✓ Colores





# Predicción de variables

Algoritmos



Predicción de variables (valores futuros)



Estudio parámetros

Modelos predictivos:

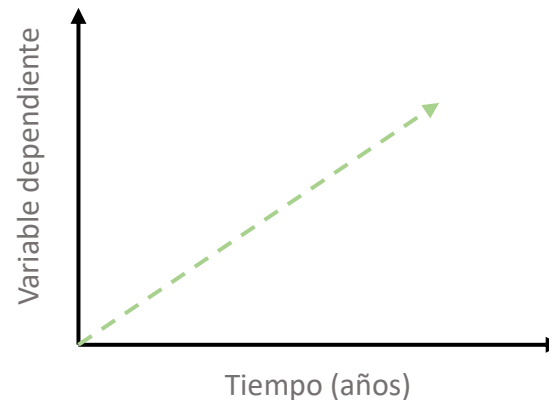
- ***Regresión lineal***
- ***Random Forest***
- ***ARIMA***



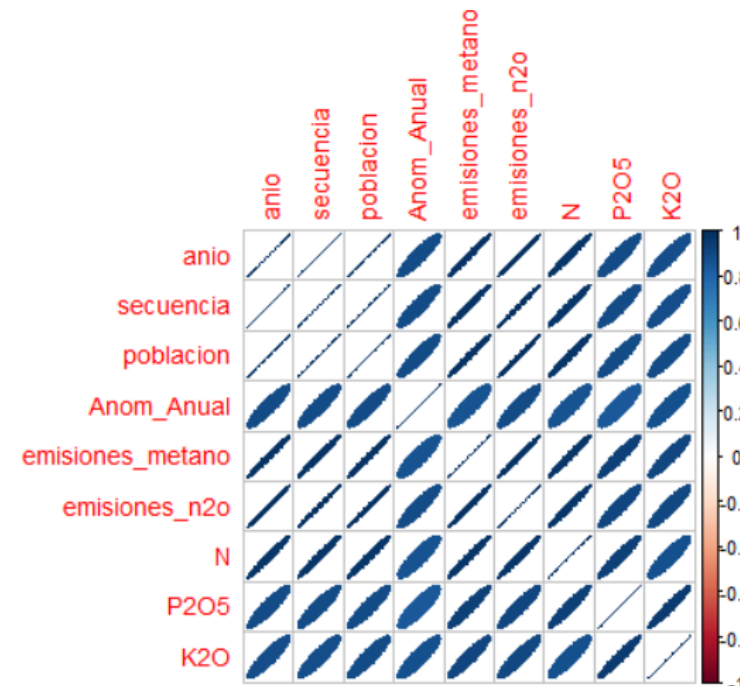
# Predicción: Regresión lineal simple

- **Ecuación**  $\rightarrow Y = \beta_0 + \beta_1 X$ 
  - ✓ Y: variable dependiente (cada una de las variables a predecir)
  - ✓ X: variable independiente (años transcurridos)
  - ✓  $\beta_0$  : ordenada al origen (intercepto).
  - ✓  $\beta_1$  : pendiente de la línea de regresión.
- **Bondad de ajuste**
  - ✓ Coeficiente de determinación (R2)
  - ✓ p-valor

**Modelo  
Aceptado**



Variable	R^2	p-valor
Uso Fertilizantes potásicos (K2O)	0,7917	4.81e-11
Temperatura	0,8032	2.16e-11
Uso Fertilizantes fosfatados (P2O5)	0,8079	1.54e-11
Uso Fertilizantes nitrogenados (N)	0,9461	2.2e-16
Emisiones Metano (CH4)	0,9629	2.2e-16
Emisiones Óxido Nitroso (N2O)	0,9834	2.2e-16
Poblacion	0,9999	2.2e-16



Regresión  
lineal



Random  
Forest



ARIMA

# Predicción: Regresión lineal múltiple

- **Ecuación**  $\rightarrow Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ 
  - ✓ Y: variable dependiente (anomalía de temperatura a predecir)
  - ✓  $X_1 \dots X_n$ : variables independientes (resto de variables)
  - ✓  $\beta_0 \dots \beta_n$ : coeficientes para ajustar el modelo.
- **Bondad de ajuste**
  - ✓ Coeficiente de determinación (R2) alto
  - ✓ p-valores altos
  - ✓ Factor de inflación de la varianza  $\rightarrow$  multicolinealidad

Call:

```
lm(formula = Anom_Anuual ~ anio + poblacion + emisiones_n2o +  
    emisiones_metano + N + P205 + K20, data = datos_temp_pobl_emis_fert)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.119300	-0.067532	0.008431	0.050296	0.147208

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-5.700e+01	5.101e+02	-0.112	0.9120
anio	2.974e-02	2.648e-01	0.112	0.9116
poblacion	-8.807e-11	3.113e-09	-0.028	0.9777
emisiones_n2o	5.589e-07	6.229e-07	0.897	0.3793
emisiones_metano	-4.698e-07	2.120e-07	-2.216	0.0373 *
N	1.502e-06	9.191e-06	0.163	0.8716
P205	3.413e-06	1.559e-05	0.219	0.8287
K20	1.616e-05	1.263e-05	1.280	0.2140

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08789 on 22 degrees of freedom  
Multiple R-squared: 0.8651, Adjusted R-squared: 0.8221  
F-statistic: 20.15 on 7 and 22 DF, p-value: 3.378e-08

```
## {r warning=FALSE}  
vif(lineal_temp_otros)
```

anio	poblacion	emisiones_n2o	emisiones_metano	N	P205	K20
20406.04279	19873.19936	87.03610	55.76511	48.60952	28.85133	20.57440

Regresión  
lineal

Random  
Forest

ARIMA



# Predicción: Regresión lineal múltiple

Call:  
lm(formula = Anom\_Anual ~ anio + emisiones\_metano, data = datos\_temp\_pobl\_emis\_fert)

Residuals:

Min	1Q	Median	3Q	Max
-0.137399	-0.066197	0.007415	0.036358	0.156764

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-6.084e+01	1.951e+01	-3.119	0.00429	**
anio	3.128e-02	1.031e-02	3.034	0.00528	**
emisiones_metano	-1.571e-07	1.579e-07	-0.995	0.32855	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0941 on 27 degrees of freedom  
Multiple R-squared: 0.8102, Adjusted R-squared: 0.7961  
F-statistic: 57.62 on 2 and 27 DF, p-value: 1.809e-10

***Modelo Descartado***

Regresión  
lineal

Random  
Forest

ARIMA





# Predicción: Regresión lineal simple

TFM - Sara Belen Ramos Gonzalez - Implementación

Search

Pages

- Home page
- Datos globales - Evolución
- Datos globales - Proporción
- Población vs Tierra Cultivada
- Precipitaciones - Evolución
- Precipitaciones - Ranking

## Trabajo Final - Máster U. en Ciencia de Datos - CULTIVOS Y CLIMATOLOGÍA. Situación Mundial.

Sara Belén Ramos González

Este informe está asociado a la realización del Trabajo final del Máster Universitario en Ciencia de Datos de la UOC, desarrollado por la alumna Sara Belén Ramos González, y titulado: **"Cultivos y Climatología. Situación Mundial"**

Se pretende permitir el análisis del comportamiento de datos relacionados con la evolución del clima y los cultivos. Concretamente, este informe muestra información sobre la temperatura, precipitaciones, emisiones de gases, usos de fertilizantes, terreno cultivado y crecimiento de la población.

### Seleccionar período y tipo de datos

Todas las páginas disponen de una segmentación de información gracias al uso de "slicers" de selección: uno para los años deseados, y otro para los tipos de datos (bien valores reales o bien valores predichos).

Además, estos "slicers" están sincronizados. Consecuentemente, al elegir un valor de año o tipo de datos, será esta selección la que se muestre en todas las páginas de manera simultánea.

1901 2050

Tipo de datos: All

### Seleccionar página a visualizar

<b>Datos Globales - Evolución</b>	<b>Datos Globales - Proporción</b>	<b>Población vs Tierra Cultivada</b>	<b>Precipitaciones - Evolución</b>	<b>Precipitaciones - Ranking</b>
Evolución gráfica sobre datos a nivel mundial de: emisiones de gases, uso de fertilizantes, anomalía de temperatura y población.	Gráficos de proporción sobre datos a nivel mundial de: emisiones de gases, uso de fertilizantes, anomalía de temperatura y población.	Evolución gráfica sobre datos a nivel mundial de: población y extensión de tierra cultivada.	Evolución gráfica sobre datos de varios países representativos sobre precipitaciones.	Información gráfica sobre el ranking de los valores de precipitaciones correspondientes a cada uno de los países estudiado.

Regresión  
lineal

Random  
Forest

ARIMA

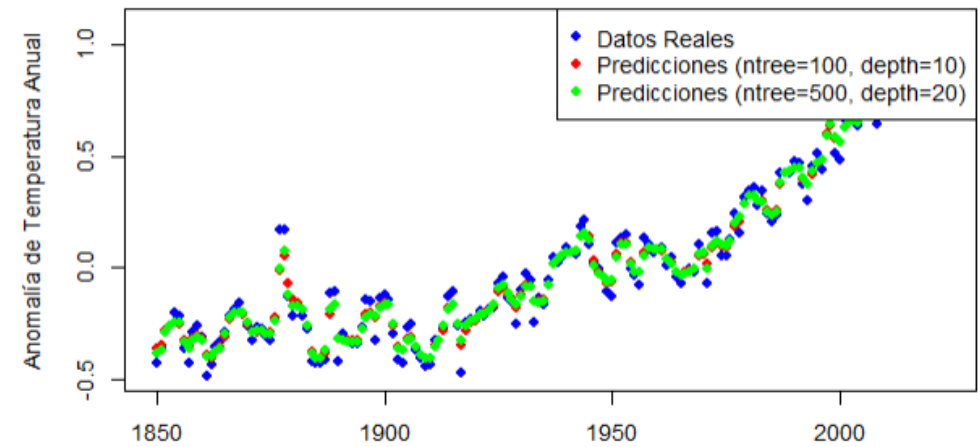


# Predicción: Random Forest

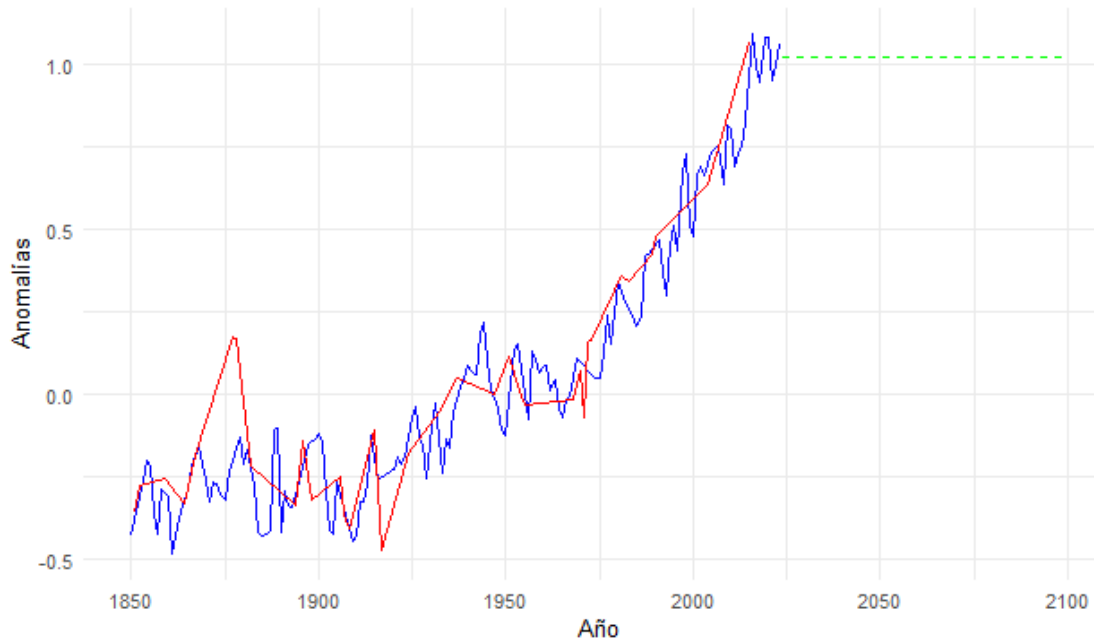
- **Predicción a largo plazo**
- **Relación no completamente lineal**
- **Parámetros**
  - ✓ MSE (error cuadrático medio) bajo.
  - ✓ Coeficiente de determinación alto.
  - ✓ No hay indicativo de sobreajuste.
  - ✓ Se han modificado los hiperparámetros.
- **Precisión de modelo**
  - ✓ Mala

**Modelo Descartado**

Random Forest - Predicciones vs. Datos Reales



Comparación de Datos de Entrenamiento, Prueba y Predicciones



Regresión  
lineal



Random  
Forest



ARIMA

# Predicción: ARIMA (AutoRegressive Integrated Moving Average)

- **Predicción series temporales**

- ✓ Autorregresivo (AR)
- ✓ Inegrado (I)
- ✓ Medias móviles (MA)

- **Estudios existentes de evolución de temperaturas**

- **Tendencia creciente en los valores**

- **Preprocesado de datos**

- ✓ Datos ordenados
- ✓ Conversión dataframe a serie temporal

```
```{r}
# Ordenar el conjunto de datos por la columna "Anio" en orden creciente
Temp_Global_Averages_1850_max <- Temp_Global_Averages_1850_max[order(Temp_Global_Averages_1850_max$Anio), ]
```

```
```{r}
# Convertir el dataframe a una serie temporal con frecuencia anual
serie_temporal <- ts(Temp_Global_Averages_1850_max$Anom_Anual, start = 1850, frequency = 1)
```



Regresión  
lineal



Random  
Forest



ARIMA

# Predicción: ARIMA

- **Bondad modelo**

- ✓ Media residuos: cercana a 0
- ✓ Varianza residuos: bajo
- ✓ Estudio residuos contra los valores ajustados sin patrones (homocedasticidad)
- ✓ Distribución de los residuos

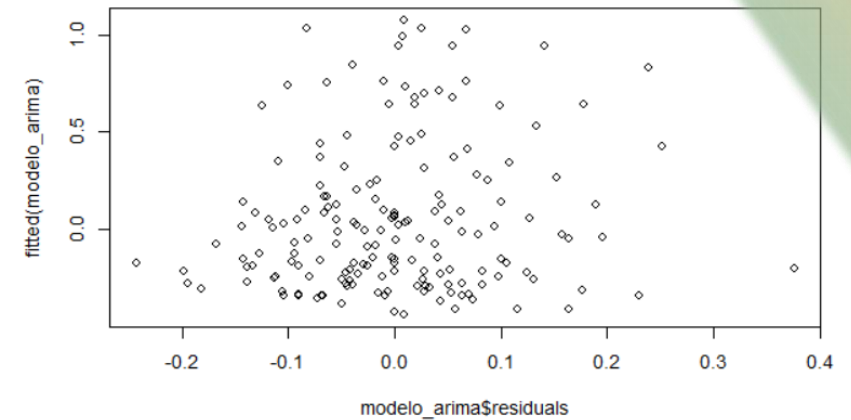
- **Equilibrio del modelo**

- ✓ Criterio de Información de Akaike (AIC): bajo
- ✓ Criterio de Información Bayesiano (BIC): bajo

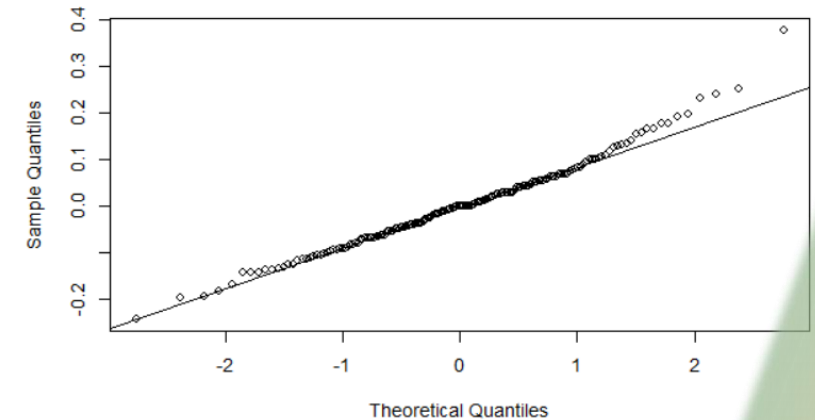
**Modelo  
Aceptado**

```
## {r}  
# Obtener los criterios de información  
aic <- AIC(modelo_arima)  
bic <- BIC(modelo_arima)  
  
# Imprimir los resultados  
cat("AIC:", aic, "\n")  
cat("BIC:", bic, "\n")
```

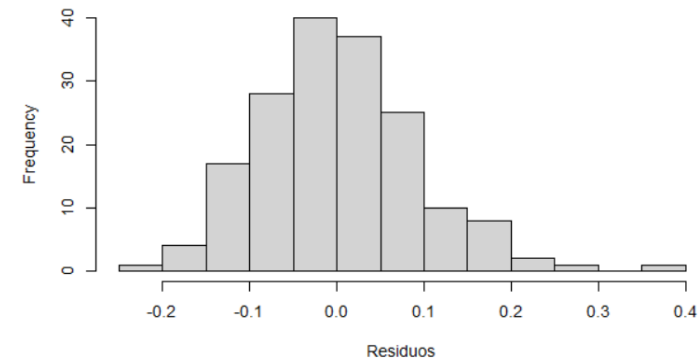
AIC: -319.509  
BIC: -303.7425



Normal Q-Q Plot



Histograma de Residuos



Regresión  
lineal



Random  
Forest



ARIMA



# Predicción: ARIMA



Regresión  
lineal



Random  
Forest

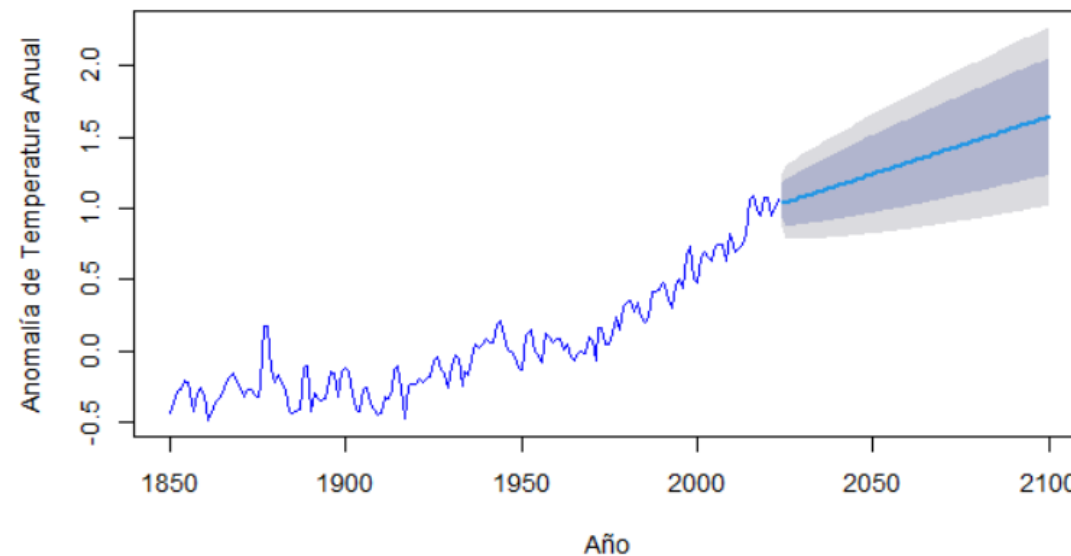


ARIMA

Anio <int>	Prediccion <dbl>
2024	1.052201
2025	1.042555
2026	1.047181
2027	1.055217
2028	1.063253
2029	1.071289
2030	1.079326
2031	1.087362
2032	1.095398
2033	1.103434

Anio <int>	Prediccion <dbl>
2034	1.111470
2035	1.119506
2036	1.127542
2037	1.135578
2038	1.143615
2039	1.151651
2040	1.159687
2041	1.167723
2042	1.175759
2043	1.183795

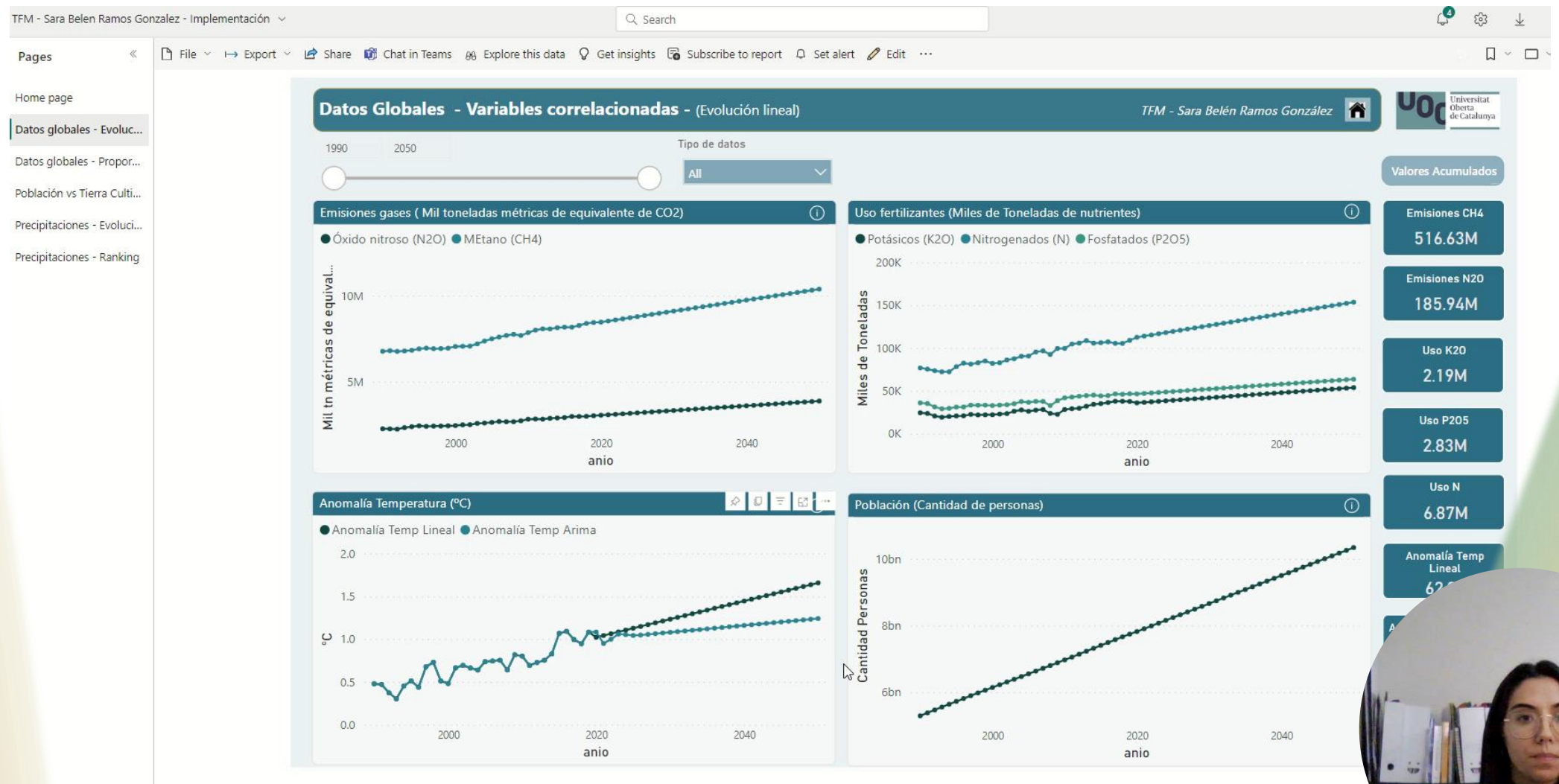
Predicciones con Modelo ARIMA



Anio <int>	Prediccion <dbl>
2044	1.191831
2045	1.199867
2046	1.207903
2047	1.215940
2048	1.223976
2049	1.232012
2050	1.240048
2051	1.248084
2052	1.256120
2053	1.264156



# Predicción: ARIMA



Regresión  
lineal



Random  
Forest



ARIMA



# Conclusiones

- ✓ **Crecimiento demográfico**
- ✓ **Intensificación agricultura**
- ✓ **Aumento de temperatura**
- ✓ **Modificación patrón lluvias**
- ✓ **Aumento de emisiones**
- ✓ **Incertidumbre climática**
- ✓ **Variabilidad hidrológica**

**Degradación Ambiental**

TFM - Sara Belen Ramos Gonzalez - Implementación

Search

Pages

- Home page
- Datos globales - Evoluci...
- Datos globales - Propor...
- Población vs Tierra Culti...
- Precipitaciones - Evoluci...
- Precipitaciones - Ranking

File Export Share Chat in Teams Explore this data Get insights Subscribe to report Edit

## Trabajo Final - Máster U. en Ciencia de Datos - CULTIVOS Y CLIMATOLOGÍA. Situación Mundial.

Sara Belén Ramos González

UOC Universitat Oberta de Catalunya

Este informe está asociado a la realización del Trabajo final del Máster Universitario en Ciencia de Datos de la UOC, desarrollado por la alumna Sara Belén Ramos González, y titulado: **"Cultivos y Climatología. Situación Mundial"**

Se pretende permitir el análisis del comportamiento de datos relacionados con la evolución del clima y los cultivos. Concretamente, este informe muestra información sobre la temperatura, precipitaciones, emisiones de gases, usos de fertilizantes, terreno cultivado y crecimiento de la población.

### Seleccionar período y tipo de datos

Todas las páginas disponen de una segmentación de información gracias al uso de "slicers" de selección: uno para los años deseados, y otro para los tipos de datos (bien valores reales o bien valores predichos).

Además, estos "slicers" están sincronizados. Consecuentemente, al elegir un valor de año o tipo de datos, será esta selección la que se muestre en todas las páginas de manera simultánea.

1901 2050

Tipo de datos

All

### Seleccionar página a visualizar

Datos Globales - Evolución	Datos Globales - Proporción	Población vs Tierra Cultivada	Precipitaciones - Evolución	Precipitaciones - Ranking
Evolución gráfica sobre datos a nivel mundial de: emisiones de gases, uso de fertilizantes, anomalía de temperatura y población.	Gráficos de proporción sobre datos a nivel mundial de: emisiones de gases, uso de fertilizantes, anomalía de temperatura y población.	Evolución gráfica sobre datos a nivel mundial de: población y extensión de tierra cultivada.	Evolución gráfica sobre datos de varios países representativos sobre precipitaciones.	Información gráfica sobre ranking de los países correspondientes a los pa...





TFM

# Cultivos y climatología. Situación Mundial.

*Sara Belén Ramos González*

Máster en Ciencia de Datos – UOC

Área 5 - Aula 1

Tutor: Rafael Luque Ocaña

PRA: Albert Solé



*Gracias por su atención.*

