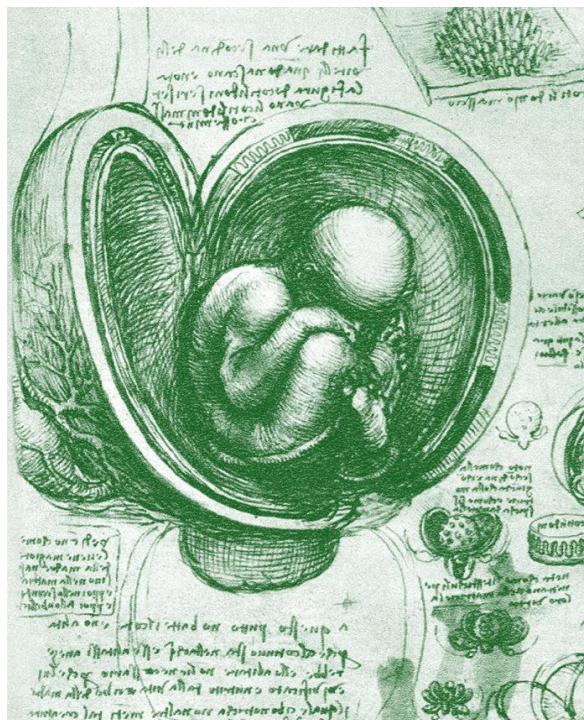


El Canvi Global a través de la Salut Materna-Filial



Alumne: Tomàs González Dantas

Grau: Biologia

Facultat: Biociències

Professora tutora: Cristina M^a Pereira Dos Santos

Tipologia: Anàlisi de dades

Data de lliurament: 25 de maig de 2026

Índex

1. Resum.....	4
1.1 Paraules clau	4
2. Antecedents	5
2.1 Contaminació atmosfèrica.....	6
2.2. Temperatura i Radiació solar	7
2.3. Aigua, Agricultura i altres	7
3. Objectius i hipòtesis	8
3.1. Objectius.....	8
3.2. Hipòtesis.....	8
4. Materials i mètodes	8
4.1. Cerca i <i>coverage</i> dels indicadors	9
4.2. <i>Outliers</i> i normalitat	11
4.3 Anàlisi de Components Principals (PCA)	12
4.4. Anàlisi multivariant.....	12
4.5. Regressions lineals múltiples.....	12
4.6. Anàlisi del polimorfisme rs1801133.....	13
5. Resultats	13
6. Discussió.....	17
7. Conclusions.....	20
8. Bibliografia	20
9. Annexos	23

Resum

La pressió que exerceix l'ambient sobre el correcte transcurs de l'embaràs pren rellevància en un món marcat pel canvi. L'augment de les temperatures i les emissions conserven una inèrcia que la fisiologia de la reproducció no pot afrontar indefinidament. Per aquesta raó, aquest treball s'emmarca en una anàlisi estadística que treballa dades agregades per països amb l'objectiu de trobar correlacions entre factors clau del Canvi Global i la Salut Materna-Filial. En el procés, es duen a terme anàlisis descriptives, multivariants i regressions múltiples amb 11 *proxies* del Canvi Global, 3 indicadors de la Salut Materna-Filial i una variable com a control del component socioeconòmic. De la discussió final, deriva un model de regressió múltiple significatiu amb una R^2 ajustada de 0,279. En aquest, el Baix Pes en Néixer s'integra com a variable dependent i la taxa de mortalitat per inaccessibilitat a aigua potable i la concentració de Matèria Particulada fina (PM2.5) representen les principals variables explicatives. A més, s'arriba a una validació qualitativa de la hipòtesi de Jablonski per a la variant rs1801133, un polimorfisme del gen *MTHFR*, relacionat amb l'alentiment del cicle del folat.

Paraules clau: Anàlisi de Components Principals (PCA), Regressió Lineal Múltiple, Matèria Particulada fina (PM2.5), Baix Pes al Néixer (BPN, LBW), rs1801133.

Antecedents

Al llarg de la història de la humanitat, els éssers humans hem anat reforçant la nostra influència sobre el Sistema Terra. Encara més, la creació i modificació de fluxos de materials que incorrem ens planteja haver entrat ja en una nova etapa geològica caracteritzada per la petjada humana. Ara bé, la consolidació d'aquesta empremta com un nou compartiment del planeta Terra implica un gran reajustament, el denominat Canvi Global¹.

Aquest fenomen no només abasta el terme Canvi Climàtic, el gran paraigua que inclou desastres naturals, anomalies tèrmiques o l'escalfament del planeta. El Canvi Global també ens parla de l'agricultura intensiva, que atempta contra la fertilitat dels sols;² de la combinació de sobreproducció i globalització, que infereix en els hàbits alimentaris;³ de l'augment de les emissions de gasos d'efecte hivernacle, la causa directa del Canvi Climàtic;⁴ dels microplàstics, metalls pesants, pesticides o altres substàncies produïdes per la indústria, que també poden acabar en organismes més enllà dels cultius⁵.

Com veiem, el concepte de Canvi Global ens pot aportar la visió que modificar els fluxos del Sistema Terra ens perjudica en moltes dimensions més enllà del Canvi Climàtic. Un terme que ha ajudat molt a conglomerar i visibilitzar aquesta qüestió és el d'exposoma; el conjunt de factors ambientals als quals ens exposem al llarg de la nostra vida⁶. No obstant això, aquest treball té la intenció de subratllar el fet que és fins i tot abans de néixer quan som més sensibles a aquest ambient.

A través del cordó umbilical ens arriben nutrients i oxigen. Tanmateix, també ens arriben totes aquelles substàncies de les quals el cos de la persona gestant encara no se n'ha pogut desfer. Aquest fet suposa una gran amenaça per al desenvolupament de l'embrió o el fetus, que encara no té les estructures necessàries per a una correcta eliminació de toxines o una homeòstasi ferma per afrontar canvis a l'ambient⁷.

En afegit, la vulnerabilitat del nadó és compartida, d'aquí que se'n parli de Salut Materna-Filial quan abordem l'efecte de l'ambient sobre la salut humana. La Salut Materna-Filial constitueix un dels components de la Salut Reproductiva; concepte que l'Organització Mundial de la Salut (OMS) defineix de la següent manera: «Un estat general de benestar físic, mental i social, i no de simple absència de malalties o dolències, en tots els aspectes relacionats amb el sistema reproductiu, les seves funcions i processos»⁸.

En aquest treball, també recollim la perspectiva de gènere que interpel·la aquesta idea; tot plegat, amb el propòsit de donar veu a nous arguments respecte a la càrrega del treball reproductiu ambiental sobre la dona i l'etiologia de les respectives malalties de l'aparell reproductor^{9,10}. A continuació, explorem

diferents dimensions en què els factors del Canvi Global i la Salut Materna-Filial intersequen:

Contaminació atmosfèrica

Diversos estudis han trobat relacions entre la Salut Reproductiva de la Dona i la qualitat de l'aire. Aquestes són tals, que es remunten fins i tot abans de la menarquia (primer episodi de menstruació). Això es donaria per l'efecte acumulat de l'exposoma sobre el reservori de fol·licles primordials; el qual no es regenera, sinó que minva en nombre de cèl·lules al llarg de la vida¹¹.

Si ens avancem fins a la qüestió de la Salut Materna-Filial, encara trobem evidència amb aquest enfocament molecular. Per exemple, s'ha vist que nivells moderats de contaminació atmosfèrica s'han associat amb canvis en la placenta, incloent-hi estrès oxidatiu i alteracions en la metilació de l'àcid desoxiribonucleic (ADN)¹². Així i tot, d'igual manera com veurem en tots els antecedents, les anàlisis que predominen són les epidemiològiques, sobretot aquelles de caràcter obstètric:

Per exemple, quant a resultats de l'embaràs, un estudi a la Xina ha demostrat que l'exposició a diòxid de sofre (SO_2), ozó (O_3), monòxid de carboni (CO) i Matèria Particulada fina (aquella menor a $2,5\ \mu\text{m}$, $\text{PM}_{2.5}$) està associada a un augment en el risc d'avortament comparat amb els controls¹³.

Si, d'altra banda, mirem en etapes més avançades de l'embaràs, trobem que per a les dones vivint a prop d'autopistes amb una alta densitat de trànsit, el risc de Baix Pes al Néixer (BPN) augmenta fins a un 2,3% (95% d'Interval de confiança (IC): 2,16–2,38) comparat a un 2,02% (95% IC: 1,90–2,12) per a aquelles dones que viuen en zones amb baixa densitat de trànsit¹⁴. Amb aquest mateix indicador de la salut del nadó i tornant a gasos específics, també s'han elaborat metaanàlisis que indiquen un increment en l'*odds ratio* (OR) del BPN davant l'exposició de diòxid de nitrogen (NO_2), CO , $\text{PM}_{2.5}$ i PM_{10} (matèria particulada menor a $10\ \mu\text{m}$). Concretament, per aquests dos últims contaminants, s'obtidrien ORs de 1,05 (95% IC: 0,99–1,12) per a la $\text{PM}_{2.5}$ a $10\ \text{mg}/\text{m}^3$ i 1,10 (95% IC: 1,05–1,15) per a la PM_{10} a $20\ \text{mg}/\text{m}^3$ ¹⁵. Així i tot, si aprofundim per exemple en els hidrocarburs policíclics aromàtics, un component de la PM, veiem que exposar-se'n durant les últimes 6 setmanes de l'embaràs té una forta associació amb el part prematur (OR: 2,74; 95% IC: 2,24–3,34)¹⁶.

Ara bé, més enllà de la $\text{PM}_{2.5}$ produïda per la crema de combustibles fòssils, també cal destacar que aquest tipus de contaminant és el que més es produeix als incendis. Respecte a això, s'ha registrat que fins i tot exposicions agudes a aquests desastres naturals cada 3 setmanes poden augmentar el risc de BPN si ajustem per característiques de la mare i el nadó¹⁷.

Temperatura i Radiació solar

Les persones gestants són especialment vulnerables a variacions tèrmiques a causa dels canvis fisiològics que experimenten, com ara un metabolisme més actiu i una relació de superfície i massa corporal reduïda¹⁸.

Entre els efectes adversos que pot tenir la temperatura sobre la Salut Materna-Filial trobem una circulació anòmla de la sang; la qual pot afectar el creixement del fetus, causar insuficiència placentària, augmentar el risc del despreniment de la mateixa placenta o limitar la producció de líquid amniòtic¹⁹. Tot plegat, contribuint a la probabilitat de patir una mort fetal:

Per exemple, un estudi de 708 dones embarassades a Texas, als Estats Units, va concloure que un augment de 10°F (5,56°C) estaria associat a un 45% de risc de patir mort fetal (OR: 1,45; 95% IC: 1,18-1,77)²⁰. Així i tot, també existeixen metaanàlisis de fins a 27 països en què es mostra que, davant l'augment d'1°C, el risc de mort fetal (OR: 1,05; 95% IC: 1,01-1,08) i part prematur augmenta (OR: 1,05; 95% IC: 1,03-1,07). Encara més, per a aquest últim indicador, el risc seria d'1,16 vegades major durant una onada de calor (OR: 1,16; 95% IC: 1,10–1,23)²¹.

Altrament, un aspecte important a destacar de l'augment de les temperatures és la sinergia que s'estableix amb els contaminants atmosfèrics. Aquesta complementaritat s'ha denotat sobretot a escala epidemiològica i en relació amb el ritme cardíac²². Així i tot, malgrat que la base fisiològica encara requereix més evidència, sí que s'ha trobat una interacció entre la radiació ultraviolada (UV) i la generació d'O₃ a prop de superfícies²³.

La inclusió de la radiació UV en la qüestió de la salut pública no només guarda una profunda relació amb els contaminants atmosfèrics; sinó que l'efecte que aquests tenen en el seu pas fins a la superfície podria inferir en el cicle del folat, clau per a un correcte desenvolupament de l'embaràs^{23,24}.

Aigua, Agricultura i altres

En els països de renda baixa i mitjana, les dones tendeixen a treballar en agricultura i altres tasques manuals fins al final de l'embaràs²⁵. Aquest esforç físic que exigeix del sector primari, sobretot en els cultius intensius, se suma als efectes que hem mencionat anteriorment, com l'estrès tèrmic i la radiació solar. No obstant això, sovint, aquest àmbit també ve acompanyat d'altres dues problemàtiques: la qualitat dels recursos hídrics i l'exposició a Disruptors Endocrins (DE).

Encara més, malgrat que no tots els DE són pesticides, com ara l'ftalat i el Bisfenol A (BPA), cal remarcar que la via per la qual els humans ens n'exposem principalment succeeix mitjançant la consumició d'aigua contaminada²⁶. A més, la seguretat al voltant aquest recurs esdevé crítica amb l'augment del nivell del

mar i els desastres naturals consegüents. Per a aquest escenari, s'espera un augment en la salinitat de l'aigua potable, la qual està associada al desenvolupament de preeclàmpsia (OR: 3,30; 95% CI: 2,00–5,51)²⁷. En afegit, quan copsem amb la dada que les dones tenen una probabilitat 14 vegades major a morir en desastres respecte als homes,²⁸ també deriva el fet que els desplaçaments dels supervivents afecten tant al projecte reproductiu com al mateix desenvolupament de l'embaràs.

Objectius i hipòtesis

Una vegada copsades les motivacions i evidències anteriors, aquest treball emergeix com un estudi ecològic en què es treballaran dades poblacionals agregades per països per tal de trobar relacions entre el Canvi Global i la Salut Materna-Filial:

Objectius

- Cercar i interpretar irregularitats en les sèries temporals d'indicadors del Canvi Global.
- Desenvolupar un model estadístic capaç d'explicar una proporció rellevant de la variabilitat en indicadors de la salut Materna-Filial a partir de *proxies* clau del Canvi Global.
- Explorar com el component socioeconòmic pot generar biaix en la interpretació de correlacions entre el Canvi Global i la Salut Materna-Filial.
- Identificar mancances en la qualitat i nombre de bases de dades públiques respecte a indicadors del Canvi Global i la Salut Reproductiva de la Dona.

Hipòtesis

- La combinació d'indicadors del Canvi Global en un mateix model estadístic és capaç de capturar una part significativa de la variabilitat en els indicadors de Salut Materna-Filial.
- Els indicadors del Canvi Global s'expressen en les anàlisis reflectint patrons típics del desenvolupament socioeconòmic.

Materials i mètodes

A continuació, repassarem l'aspecte més tècnic de cadascuna de les fases del treball. Concretament, descriurem els mètodes de cerca i estructuració de les dades, les eines emprades per a treballar-les i els criteris estadístics aplicats durant la seva anàlisi.

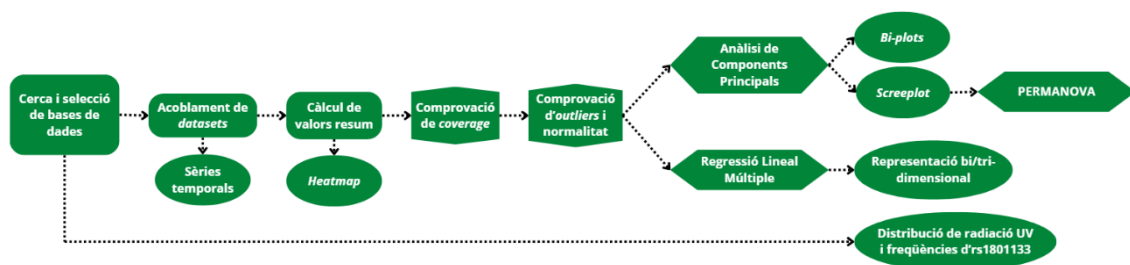


Figura 1. Diagrama de la metodologia emprada. Rectangles = estructuració de les dades; El·lipses = representacions gràfiques; hexàgons = criteris i models estadístics.

Cerca i coverage dels indicadors

La selecció d'indicadors s'ha fonamentat en un recull representatiu d'aquells elements més rellevants segons la bibliografia. Aquesta compilació inclou diferents indicadors de la Salut Materna-Filial, *proxies* de diferents compartiments del Sistema Terra i el Canvi Global i almenys una variable suplementària o covariable control de l'estatus socioeconòmic. Gairebé totes les dades han estat extretes de fonts institucionals de gran renom, com ara l'OMS, el *World Bank* o el *United Nations Children's Fund* (UNICEF); en aquest sentit, només un indicador ha estat extret d'un article, concretament de la revista *Springer Nature*. [Taula 1]

Quant a la naturalesa de les dades, ens trobem que, per a alguns d'aquests indicadors, només disposem d'una mesura puntual en el temps (variables *snapshot*). Així i tot, una gran proporció dels indicadors s'expressa en forma de sèries temporals. És tant així que, donat el desfasament entre *datasets* que deriva d'aquesta qüestió, s'ha acotat una finestra temporal que optimitzi el grau de superposició temporal. Concretament, el lapse escollit ha estat 2000-2020 i el subsegüent encavalcament de les bases de dades s'ha dut a terme mitjançant un script [Script 1; Annexos].

Un cop tots els indicadors s'han acoblat mitjançant el codi ISO 3166-1 alfa-3 de cadascun dels països i han estat automàticament classificats per les regions que estableix l'OMS (Àfrica – AFR, Amèriques – AMR, Mediterrani Oriental – EMR, Europa – EUR, Sud-Est Asiàtic – SEAR, Pacífic Occidental - WPR), s'ha procedit amb un càlcul que extreu valors resum de les sèries temporals. Respecte a això, per a cada indicador i període, s'han calculat les mitjanes triennals i totals i el canvi percentual entre la dada més antiga i la més recent. Així i tot, hi ha hagut alguns indicadors que per la variabilitat o signe de les seves dades, s'ha realitzat una suma de valors absoluts que permeti la seva interpretació.

Tots aquests càlculs s'han realitzat mitjançant un script d'*RStudio*²⁹ que, a més, extreu el *coverage* de cadascun dels indicadors. [Script 2-3, Output 1-2; Annexos] Ha estat, doncs, amb aquest darrer valor que s'han filtrat i eliminat aquells indicadors amb una cobertura de països prou llunyana de la mitjana

(≈80% excloent variables *snapshot*) per a interferir en l'anàlisi. En aquest cas, s'ha aplicat un criteri de *coverage* < 60%, tant per als Valors Mitjans com per als Canvis Percentuals. A la Taula 1 s'expressen tots els indicadors esmentats, el seu codi, les seves unitats i com han estat processats:

Taula 1. Taula resum amb tots els indicadors considerats, dividits per blocs temàtics. Els indicadors marcats en vermell han estat eliminats pel criteri de coverage. Sigles de les fonts de dades: Our World in Data (OWID); World Bank (WB); World Health Organisation (WHO); Food and Agriculture Organisation (FAO); International Labour Organisation (ILO); Global Health Observatory (GHO); United Nations Development Project (UNDP).

Tipus	Codi	Institució	Descripció de l'indicador	Unitats	Càlcul
Contaminants atmosfèrics	CO2	OWID	Emissions de CO ₂ per càpita derivades de la crema de combustibles fòssils i la indústria, expressades en tones de CO ₂ per persona i any.	Tones de CO ₂ per càpita (t CO ₂ /persona/any)	Canvi percentual i mitjanes
	PM2.5	WB / WHO	Concentració mitjana anual de PM2.5 a l'aire ambient, ponderada per la població exposada.	Micrograms per metre cúbic (µg/m³)	Canvi percentual i mitjanes
Usos del sòl	Pesticides	OWID / FAO	Ús de pesticides (insecticides, herbicides i fungicides) per unitat de superfície de terres de conreu.	Quilograms per hectàrea de terres de conreu (kg/ha)	Canvi percentual i mitjanes
	Occupation	WB / ILO	Percentatge de dones empleades en el sector agrícola respecte al total de dones ocupades.	Percentatge de l'ocupació femenina total (%)	Canvi percentual i mitjanes
Aigua	HidroM	WHO - GHO	Taxa de mortalitat atribuïda a l'exposició a serveis d'aigua, sanejament i higiene (WASH) no segurs (SDG 3.9.2).	Morts per 100.000 habitants	Snapshot
	Hidro	WHO - GHO	Percentatge de la població que utilitza serveis d'aigua potable gestionats de manera segura, d'acord amb els criteris de l'OMS/UNICEF JMP.	Percentatge de la població (%)	Canvi percentual i mitjanes
	Poisoning	WHO - GHO	Taxa de mortalitat atribuïda a enverinaments no intencionals (per exposició a substàncies nocives, excloent intoxicacions per drogues).	Morts per 100.000 habitants	Canvi percentual i mitjanes
Malalties	Malaria	WHO - GHO	Incidència de malària estimada per 1.000 persones en risc d'infecció a l'any.	Casos per 1.000 persones en risc	Canvi percentual i mitjanes
Desastres naturals	Affecteds	OWID / EM-DAT	Nombre total de persones afectades per desastres naturals (inclou morts, ferits, desplaçats i necessitats d'assistència d'emergència).	Nombre de persones afectades (n)	Suma de valors absoluts
Salut Materna-Filial	LBW	UNICEF	Nombre estimat de nadons nascuts amb un pes inferior a 2.500 grams (baix pes en néixer).	Nombre de nadons (n per any)	Canvi percentual i mitjanes
	HDP	Springer / BMC Pregnancy and Childbirth ³⁰	Prevalença de trastorns hipertensius de l'embaràs (preeclàmpsia, eclàmpsia i hipertensió gestacional) a partir d'una revisió sistemàtica i metaanàlisi global.	Percentatge de gestants afectades (%)	Snapshot
	Stillbirth	WHO - GHO	Taxa de mortinadons (nascuts morts a partir de les 28 setmanes de gestació) per 1.000 naixements totals.	Mortinadons per 1.000 naixements totals	Canvi percentual i mitjanes
	Preterm	WHO - GHO	Taxa de naixements prematurs (abans de les 37 setmanes de gestació) per 100 naixements vius.	Naixements prematurs per 100 naixements vius (%)	Snapshot
Temperatura i radiació	Anomaly	OWID / Berkeley Earth	Anomalia de temperatura superficial anual respecte a la mitjana preindustrial de referència (1850–1900), calculada a escala nacional.	Graus Celsius (°C) respecte al període de referència	Suma de valors absoluts
	Temp	OWID / Berkeley Earth - ERA5	Temperatura mitjana anual de la superfície terrestre per país, derivada de reanàlisis climàtiques i registres d'estacions meteorològiques.	Graus Celsius (°C)	Suma de valors absoluts

	UV	WHO – GHO	Nivell d'irradiació ultraviolada solar a la superfície terrestre, expressat com a índex UV anual mitjà ponderat per la població.	Índex UV (UV-índex, adimensional)	Snapshot
Salut alimentària	BMI	WHO – GHO	Prevalença de sobrepès en adults (IMC ≥ 25 kg/m ²), estimació estandaritzada per edat, ambdós sexes combinats.	Percentatge de la població adulta (%)	Canvi percentual i mitjanes
Socio-economia	HDI	UNDP	Índex de Desenvolupament Humà compost per esperança de vida, educació i renda nacional bruta per càpita, en una escala de 0 a 1.	Índex adimensional (0–1)	Canvi percentual i mitjanes

Outliers i normalitat

Un cop seleccionats els indicadors i la finestra temporal en la qual s'estableix el treball, primer s'ha comprovat la presència d'*outliers* que poguessin entorpir l'anàlisi. Per a aquesta tasca, s'han seguit dues vies: Per una banda, s'han generat dos *heatmaps* amb *RStudio*, un per a les mitjanes de les sèries temporals 2000-2020 i un altre per al Canvi Percentual dels indicadors pertinents. [Script 4-5, Figures I-II; Annexos] D'altra banda, a *Jamovi*³¹, s'han identificat els *outliers* a partir de *boxplots* i *QQ-plots* de cadascun dels indicadors, segregats per regió.

Un cop identificats, el criteri de validació final per als *outliers* s'ha basat en els següents aspectes: En primer lloc, s'han considerat dins de l'anàlisi aquells *outliers* sorgits pel comportament del càlcul resum; ja que en aquells països amb xifres inicials realment petites, qualsevol increment pot suposar un canvi percentual accentuat. Alhora, també s'han acceptat aquells països considerats *outliers* dins de la seva regió, però que quadrarien en la divisió territorial fronterera. Per acabar, tenim aquells casos relacionats amb el caràcter industrial de certs països, en què les xifres accentuades estarien justificades sota una predominança de les finances respecte als sistemes de producció o viceversa, per exemple.

En definitiva, els *outliers* eliminats han estat aquells que pel seu caràcter insular encara estan procés de millora estadística, els que tenen una població molt reduïda o una molt baixa transparència de dades. En total, aquests *outliers* representarien un 4,2% i un 1,9% del total de països en la versió original dels *datasets* de Valors Mitjans i Canvis Percentuals, respectivament.

Sigui com sigui, tot just després d'aquest filtre, s'ha comprovat la normalitat dels residus dels valors resums per a esbrinar si continuar l'anàlisi amb proves paramètriques o proves estadístiques de distribució lliure. Per a testar-ho, s'ha executat la prova Shapiro Wilk (normalitat a p-valor $> 0,05$) abans i després d'una transformació de $\log_{10}$. No obstant això, com la majoria dels indicadors no mostren p-valors ni histogrames de densitat propis de distribucions normals, s'ha procedit amb una anàlisi no paramètrica.

Anàlisi de Components Principals (PCA)

Tenint en compte que després d'aplicar els anteriors filtres es compta amb 11 indicadors del Canvi Global [Taula 1], es decideix emprendre l'objectiu del model multivariant iniciant l'anàlisi amb un PCA. Aquesta fase s'ha dut a terme per dues vies:

Per una banda, mitjançant *Jamovi*, s'ha configurat el PCA incloent els indicadors del Canvi Global com a variables actives i l'HDI i els 3 indicadors de la Salut Materna-Filial validats (Stillbirth, HDP i LBW) com a variables suplementàries o de control. D'aquesta configuració, s'extreuen les dimensions amb les seves respectives coordenades.

D'altra banda, el treball realitzat a *RStudio* s'ha focalitzat en la representació dels resultats del PCA. [Script 6-7; Annexos] Els *screeplots* generats han permès seleccionar les dimensions sota els criteris d'assolir aproximadament el 60% de la variància acumulada tot respectant el criteri de Kaiser (exclusió de dimensions amb *Eigenvalue* <1). És, doncs, un cop seleccionades les dimensions, que s'han realitzat tots els *bi-plots* respectius possibles. En aquesta última tasca, no només aconseguim representar la regió a la qual pertany cadascun dels països, sinó que, a més, es realitza una clusterització mitjançant k-mitjanes.

Anàlisi multivariant

Un cop seleccionades les dimensions, mitjançant *Jamovi*, es comprova si existeix normalitat multivariant que validés una prova paramètrica com l'Anàlisi Multivariat de Covariància (MANCOVA). No obstant això, en no donar-se una distribució normal dels residus, es procedeix amb un Anàlisi Permutacional Multivariat de la Variància (PERMANOVA), generat mitjançant un script d'*RStudio* [Script 8; Annexos]. Quant a aquesta última prova, es construeixen els següents 4 models per a intentar explicar la variància de les 3 variables dependents de la Salut Materna-Filial:

- Només les dimensions seleccionades del PCA (Dims)
- Dims + HDI
- Dims + Regió
- Dims + HDI + Regió

Regressions lineals múltiples

Un cop dut a terme l'anàlisi multivariant, es procedeix amb una regressió lineal múltiple per a cadascun dels tres indicadors de la Salut Materna-Filial. De nou, amb *Jamovi*, s'han eliminat els *outliers* dels residus i comprovat la normalitat de la seva distribució (Shapiro Wilk, p-value > 0,05). Posteriorment, s'ha efectuat un control de la col·linealitat mitjançant el Variation Inflation Factor (VIF) amb el criteri d'inclusió en l'anàlisi d'acord amb un VIF < 5. Per acabar, aquells models

amb major potencial han estat optimitzats amb la prova Akaike, acceptant l'exclusió de variables sota el criteri d'una millora de 2 punts sobre aquest índex.

Anàlisi del polimorfisme rs1801133

Finalment, com un control afegit per a l'anàlisi, es procedeix a considerar un component genètic mencionat a la bibliografia. En concret, s'ha abordat la selecció balancejadora sobre el polimorfisme rs1801133, el qual alenteix el metabolisme del folat a nivell de la proteïna MTHFR. Aquesta selecció s'ha testat mitjançant el visualitzador *PopHuman*; més específicament, emprant el *track D* de Tajima i cercant a poblacions de diferents continents: CEU, ACB, DEB, BEB, CDX i CLM.

Posteriorment, s'ha decidit fer una valoració qualitativa de la distribució de freqüències al·lèliques d'aquest polimorfisme tot representat simultàniament la radiació UV registrada a cada país. Per a dur-ho a terme, s'han obtingut les freqüències al·lèliques de totes les mostres no diaspòriques del projecte 1000 genomes i s'han indexat gràfics de sectors a les coordenades respectives d'un mapamundi. [Script 9; Annexos]

Resultats

Pel que fa a la representació de les sèries temporals, copsem diferents tipus de patrons. [Figura III; Annexos] Per una banda, trobem aquells indicadors en què totes les regions segueixen una tendència positiva, com succeeix amb l'*HDI* i *BMI*. D'altra banda, tenim aquelles variables en què totes les regions expressen xifres tendint a la baixa, alguns exemples serien *Occupation* i *Stillbirth*. Després, tenim aquelles sèries temporals en què, malgrat haver-hi tendències dispars entre els continents, tots evolucionen a un ritme propi i ordenat o es troben en xifres constants en el temps. Exemples d'aquest últim tipus serien els casos per al *CO₂*, els Pesticides i *LBW*. No obstant això, també reconeixem un últim tipus d'indicador, aquell en què les regions segueixen un patró irregular al llarg del temps. [Figura 2] Per exemple, en el cas de la sèrie temporal de *Poisoning*, es denota que l'Àfrica arriba a un màxim molt subtil l'any 2013. Així i tot, aquest fenomen es veu de manera molt més marcada en el gràfic de PM2.5; on, a partir del 2010, les mesures comencen a expressar prou més variabilitat.

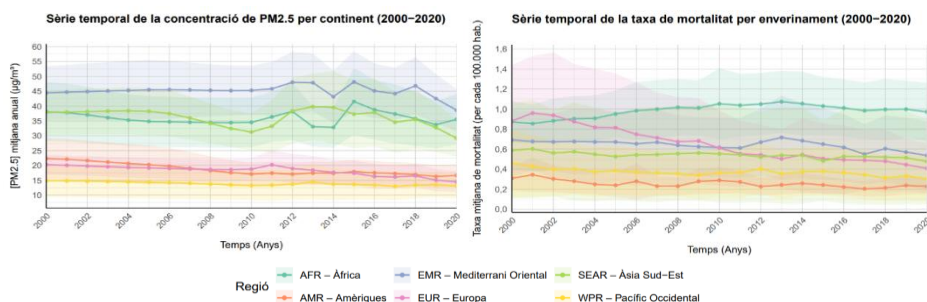


Figura 2. Sèries temporals (2000-2020) dels indicadors PM2.5 (esquerra) i Poisoning (dreta).

Deixant de banda l'anàlisi descriptiva i passant a la multivariant, veiem que els *scree-plots* deixen de complir el criteri de Kaiser i assoleixen el 60% de la variància acumulada un cop considerades 3 i 4 dimensions per als PCAs de Valors Mitjans i Canvis Percentuals, respectivament. [Figura 3]

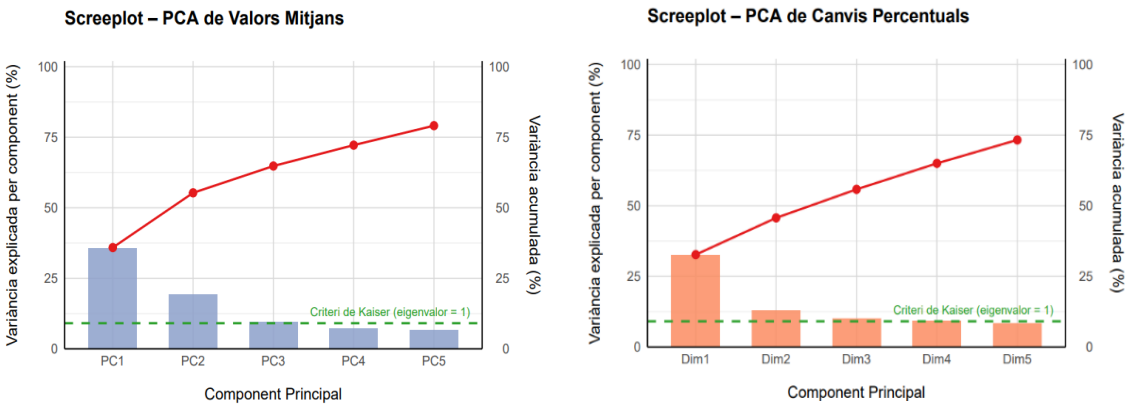


Figura 3. Scree-plots dels PCAs de Valors Mitjans (esquerra) i Canvis Percentuals (dreta) mostrant el perfil de variància explicada al llarg de les primeres 5 dimensions amb major poder explicatiu.

Davant aquests resultats dels *scree-plots*, es construeixen tots els *bi-plots* possibles per a cada cas i els respectius PERMANOVAs. A continuació, s'exposen aquells *bi-plots* més interessants de cara a la discussió dels resultats i una taula resum del PERMANOVA. [Figures 4-5; Taula 3]

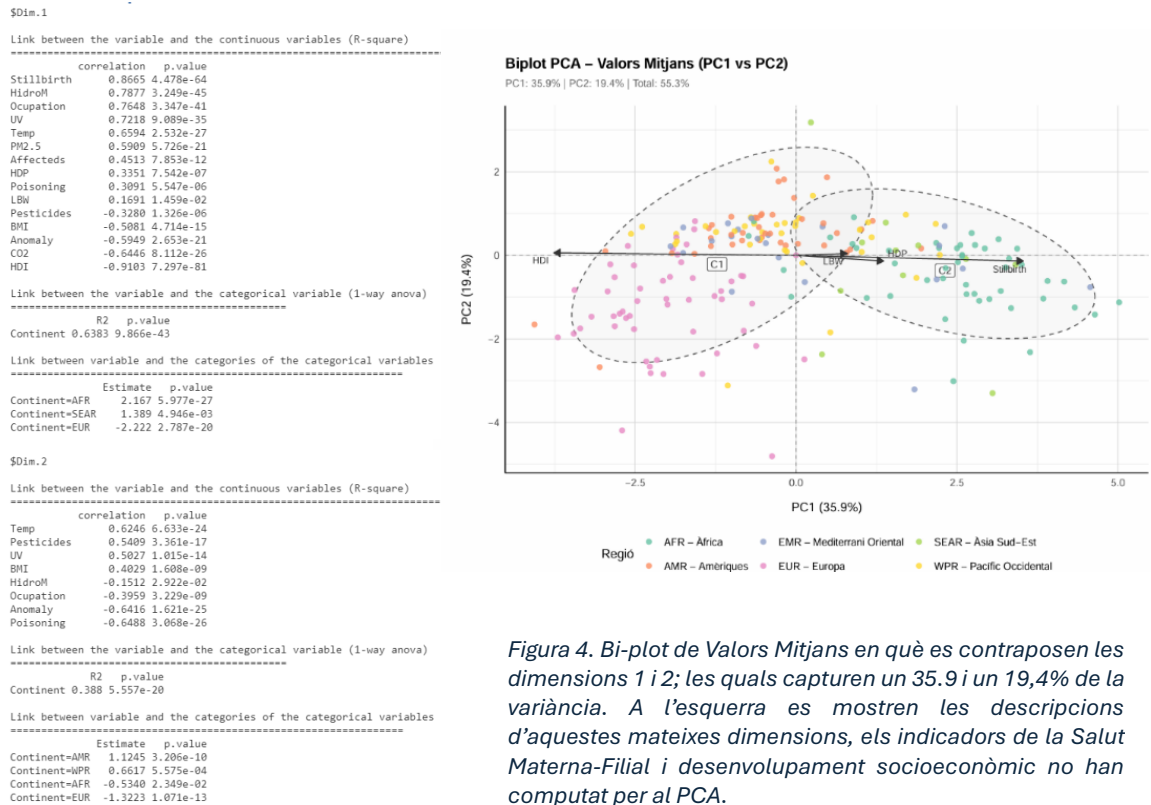


Figura 4. Bi-plot de Valors Mitjans en què es contraposen les dimensions 1 i 2; les quals capturen un 35.9 i un 19.4% de la variància. A l'esquerra es mostren les descripcions d'aquestes mateixes dimensions, els indicadors de la Salut Materna-Filial i desenvolupament socioeconòmic no han computat per al PCA.

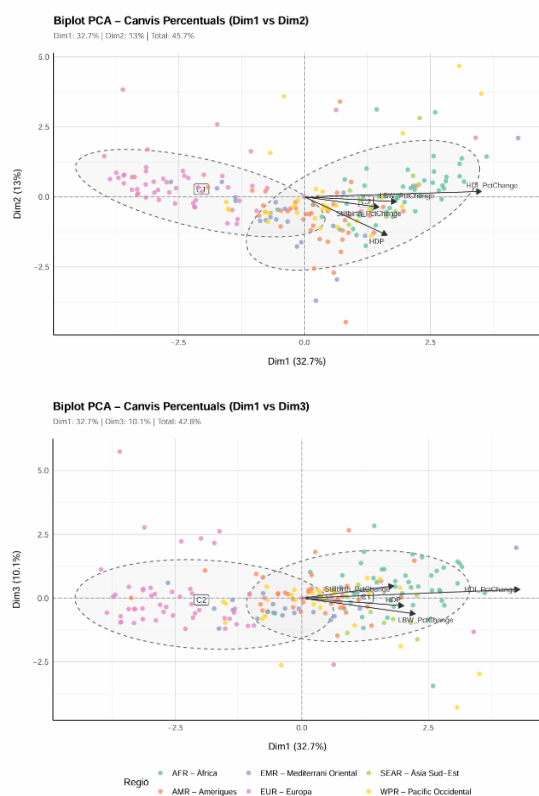


Figura 5. Bi-plots de Canvis Percentuals en què es contraposa la dimensió 1 amb la 2 i la 3, respectivament. A la dreta es mostren les descripcions d'aquestes mateixes dimensions.

\$Dim.1	
Link between the variable and the continuous variables (R-square)	
	correlation p.value
UV	0.8527 1.349e-58
Temp_AbsSum	0.8289 1.329e-52
BMI_PctChange	0.6647 3.017e-27
HidroM	0.6619 5.835e-27
HDI_PctChange	0.5136 4.638e-15
PM2.5_PctChange	0.4782 5.373e-13
Affecteds_AbsSum	0.4780 5.536e-13
CO2_PctChange	0.3992 3.627e-09
HDP	0.2472 3.767e-04
Occupation_PctChange	0.2126 2.323e-03
Pesticides_PctChange	0.2043 3.462e-03
LBW_PctChange	0.1639 1.948e-02
Anomaly_AbsSum	-0.7712 2.688e-41
Link between the variable and the categorical variable (1-way anova)	
	R2 p.value
Continent	0.7225 6.646e-53
\$Dim.2	
Link between the variable and the continuous variables (R-square)	
	correlation p.value
CO2_PctChange	0.5562 6.991e-18
Poisoning_PctChange	0.4077 1.564e-09
BMI_PctChange	0.4003 3.266e-09
Pesticides_PctChange	0.3805 2.144e-08
Anomaly_AbsSum	0.3176 3.879e-06
HDI_PctChange	0.2980 1.577e-05
HidroM	0.2705 9.514e-05
Affecteds_AbsSum	0.2696 1.003e-04
UV	-0.2408 5.375e-04
LBW_PctChange	-0.2453 4.194e-04
Stillbirth_PctChange	-0.2705 9.511e-05
Temp_AbsSum	-0.3891 9.589e-09
Occupation_PctChange	-0.4312 1.346e-10
Link between the variable and the categorical variable (1-way anova)	
	R2 p.value
Continent	0.1882 7.865e-08
\$Dim.3	
Link between the variable and the continuous variables (R-square)	
	correlation p.value
Poisoning_PctChange	0.7798 9.275e-43
HidroM	0.3701 5.480e-08
Stillbirth_PctChange	0.2285 1.042e-03
Affecteds_AbsSum	0.1436 4.097e-02
BMI_PctChange	-0.1619 2.101e-02
Pesticides_PctChange	-0.3209 3.035e-06
CO2_PctChange	-0.4428 3.711e-11
Link between the variable and the categorical variable (1-way anova)	
	R2 p.value
Continent	0.1098 0.0003248

Taula 2. Resultats del PERMANOVA per a Valors Mitjans i Canvis Percentuals. S'hi especifica el nombre de països, la R^2 dels components a cadascun dels models i la R^2 ajustada (R^2 aj.) del model complet. En verd, s'han marcat les correlacions significatives, ns = no significatiu.

	Valors mitjans · N=151				Canvi percentual (2000–2020) · N=148			
	Dim sol	Dim+HDI	Dim+Cont	Dim+HDI+Regió	Dim sol	Dim+HDI	Dim+Cont	Dim+HDI+Regió
Dim1	0,087 p<0,001	0,072 p<0,001	0,013 ns	0,025 p=0,051	0,041 p=0,012	0,008 ns	0,000 ns	0,000 ns
Dim2	0,003 ns	0,011 ns	0,008 ns	0,014 ns	0,013 ns	0,015 ns	0,014 ns	0,014 ns
Dim3	0,010 ns	0,009 ns	0,009 ns	0,008 ns	0,000 ns	0,000 ns	0,000 ns	0,000 ns
Dim4	—	—	—	—	0,000 ns	0,001 ns	0,000 ns	0,000 ns
HDI	—	0,031 p=0,026	—	0,013 ns	—	0,050 p=0,017	—	0,018 p=0,092
Regió	—	—	0,061 p=0,114	0,042 ns	—	—	0,108 p=0,021	0,076 p=0,064
R^2 aj.	0,081	0,107	0,112	0,119	0,027	0,073	0,107	0,120

Pel que fa als models de regressió lineal múltiple, només 2 d'ells compleixen normalitat de residus després de la correcció per col·linealitat i refinament pel mètode Akaike:

En primer lloc, tenim una correlació amb una R^2 ajustada de 0,128 amb *Stillbirth* com a variable dependent [Figura 6]. No obstant això, si ens fixem en els coeficients i la seva significança, no trobarem cap contribució consolidada.

No obstant això, quan abordem el canvi percentual del BPN, malgrat s'ha hagut descartar Qatar per l'*outlier* que representa el seu residu, trobem una correlació moderada (R^2 ajustada = 0,279) amb un model relativament simple. A continuació, es representen les correlacions de BPN amb els dos indicadors més rellevants. [Figura 7]

Mesures de Ajustament del Model					A	Mesures de Ajustament del Model					B
Model	R	R ²	R ² ajustada	AIC (Akaike)		Model	R	R ²	R ² ajustada	AIC (Akaike)	
1	0.421	0.177	0.128	1295		1	0.571	0.326	0.279	1159	
Nota. Models estimated using sample size of N=162						Nota. Models estimated using sample size of N=126					

Coeficients del model - Stillbirth_PctChange					A	Coeficients del model - LBW_PctChange					B
Predictor	Estimador	EE	t	p		Predictor	Estimador	EE	t	p	
Constant	-34.71785	7.78289	-4.461	1.58e-5		Constant	5.0968	7.22242	0.706	.48179	
CO2_PctChange	-0.02093	0.00850	-2.462	.0149		PM2.5_PctChange	0.5446	0.11165	4.878	3.42e-6	
Temp_AbsSum	0.01768	0.00978	1.808	.0726		HidroM	0.3253	0.14664	2.219	.02845	
Affecteds_AbsSum	2.78e-6	2.19e-5	0.127	.8992		CO2_PctChange	-0.0648	0.02471	-2.624	.00986	
Anomaly_AbsSum	-0.15070	0.56617	-0.266	.7905		HDI_PctChange	0.1094	0.23490	0.466	.64222	
Poisoning_PctChange	0.00355	0.01544	0.230	.8183		Affecteds_AbsSum	-1.01e-4	4.50e-5	-2.253	.02613	
BMI_PctChange	-0.03282	0.02615	-1.255	.2114		Pesticides_PctChange	-8.79e-4	0.00148	-0.595	.55275	
PM2.5_PctChange	0.01252	0.05707	0.219	.8267		Occupation_PctChange	-0.0491	0.04622	-1.062	.29049	
HidroM	0.11612	0.06913	1.680	.0951		Temp_AbsSum	0.0176	0.01459	1.209	.22912	
HDI_PctChange	-0.18083	0.10237	-1.766	.0793							

Figura 6. Variables dependents de les regressions lineals múltiples després de les correccions de col·linealitat i Akaike. Model per a Stillbirth (4A). Model per a BPN (4B).

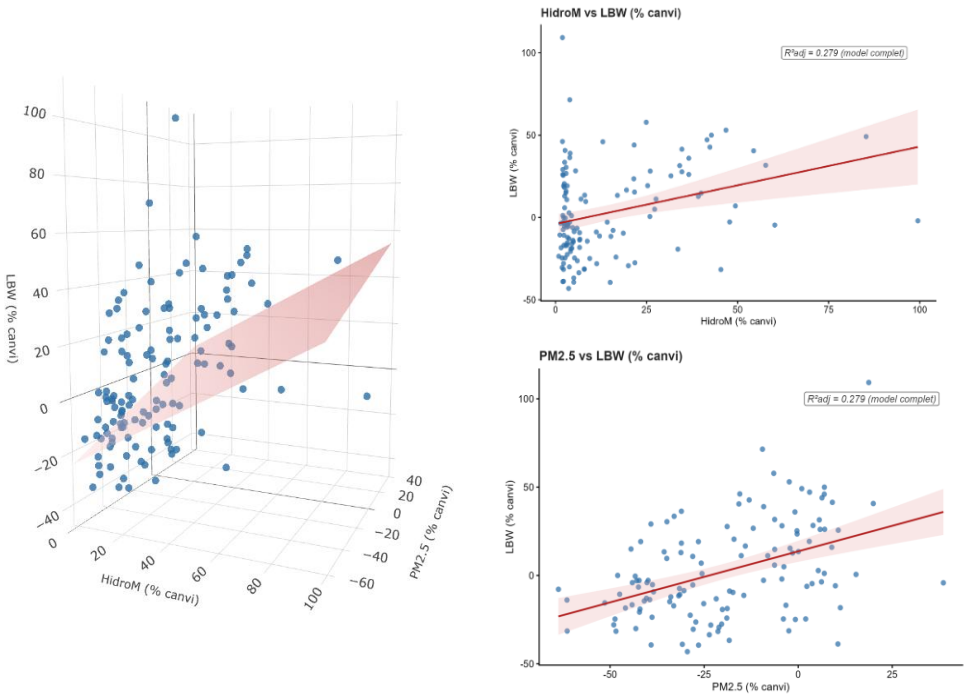


Figura 7. A la dreta: representació bidimensional de la correlació entre BPN (LBW) i HidroM i PM2.5, individualment. A l'esquerra: representació tridimensional de la correlació entre LBW, HidroM i PM2.5, simultàniament.

Per a finalitzar, quan cerquem el gen *MTHFR* al *PopHuman*, els *tracks* de *Tajima's D* no marquen cap xifra per sobre del 0,2; és a dir, no es denota cap indicati de selecció. No obstant això, si representem la distribució de les freqüències al·lèliques respecte a la de la radiació UV, s'albira una certa relació. [Figura 8]

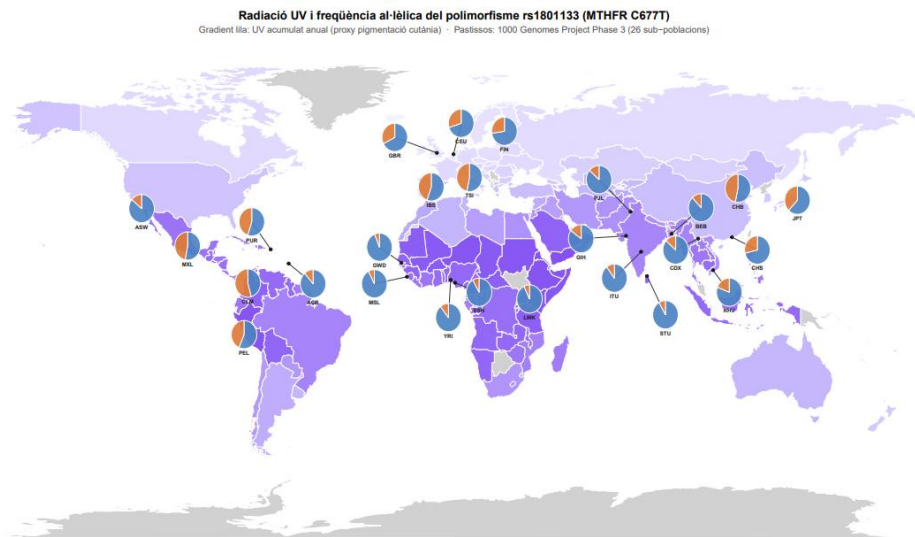


Figura 8. Representació del gradient de radiació UV i la freqüència al·lèlica del polimorfisme rs1801133 com a gràfics de sectors indexats al lloc de mostreig. Les freqüències al·lèliques referents als nucleòtids A (variant que alenteix el cicle del folat) i G corresponen, respectivament, al taronja i al blau.

Discussió

Com hem comentat en els resultats, les sèries temporals dels indicadors mostren diferents tipus de patrons. No obstant això, la gran majoria s'expressen d'acord amb les transicions demogràfiques, socioeconòmiques i energètiques lligades al Canvi Global. Per exemple, quan abordem la concentració de PM_{2.5} i CO₂, veiem que les regions on predominen les economies avançades presenten perfils baixos o tendències cap a la reducció d'aquests contaminants; mentre que les regions amb una indústria en ple desenvolupament augmenten les concentracions o se situen per sobre a escala de valors absoluts. Ara bé, a regions com l'Àfrica, ens trobem que malgrat situar-nos en mínims de CO₂, aquesta regió manté una de les concentracions de PM_{2.5} més elevades.

La compensació entre aquests dos tipus de contaminants atmosfèrics també l'hem capturada a la Dimensió 1 del PCA de Valors Mitjans. [Figura 4] És, doncs, a partir d'aquest *bi-plot*, que les projeccions dels Eigenvectors suggereixen i reforcen l'argument que les xifres elevades de CO₂ són característiques de regions amb un fort component d'indústria i transport; mentre que el PM_{2.5} estaria més associat a fonts d'energia tradicionals com ara la llenya i el carbó, la principal font d'energia als països de l'Àfrica^{32,33}.

Deixant això a banda, també caldria destacar altres dues propietats de les sèries temporals: Per una banda, les irregularitats que trobem en el perfil temporal del PM2.5; probablement causades per la crisi financera del 2008³⁴. D'altra banda, el fet que l'augment gradual en l'HDI ve acompanyat d'una baixada suau en el *LBW* i l'*Stillbirth*, alhora que un augment en els indicadors *BMI* i Hidro, mostrant una millora en la salut pública i la seguretat dels recursos hídrics.

Aprofundint ara en l'anàlisi multivariant, veiem que, tant en el PCA de Valors Mitjans com en el de Canvis Percentuals, es denota una forta diferenciació entre les economies emergents de l'Àfrica i els països d'Europa. Concretament, en el cas dels Valors Mitjans de l'Àfrica, trobem que els contaminants atmosfèrics s'agreguen amb les altes i estables temperatures, la radiació UV i l'ocupació agrària femenina; tot plegat, en la mateixa direcció en què ho faria l'indicador de mort fetal.

Ara bé, un altre tret a destacar d'aquesta secció és que, en tots els *bi-plots*, els clústers associats a l'Àfrica i a Europa se superposen generant una zona en la qual s'ubiquen una gran quantitat d'economies de transició. Aquesta clusterització podria ser una de les raons per les quals denotem que l'*HDI* i la classificació per regions absorbeixen part important de la significança de les dimensions un cop s'integren en un model PERMANOVA. Així i tot, ens trobem que en el model de Dims + HDI per als Valors Mitjans, la Dimensió 1 i l'indicador del desenvolupament guarden significança simultàniament. Encara més, la correlació que marcaria la Dimensió 1 seria superior a la de l'HDI, la qual cosa donaria marge per a una intervenció real dels factors ambientals sobre els indicadors de la Salut Materna-Filial.

No obstant això, donats els resultats poc satisfactoris, es va decidir procedir amb una regressió lineal múltiple. En aquest cas, comentarem únicament la regressió que ha complert el criteri de normalitat en la distribució dels residus, aquella amb el *LBW* com a variable dependent. Per a aquesta regressió, els indicadors amb més contribució són el Canvi Percentual de PM2.5 i la variable *snapshot* HidroM. Concretament, aquests indicadors resulten interessants; ja que recordem que, malgrat que la regió d'Àfrica predomina en concentració de PM2.5, el canvi que mostra en el perfil de la sèrie temporal és mínim. Encara més, la tendència per a totes les regions apunta cap a una reducció d'aquestes concentracions.

Per tal d'aprofundir, doncs, en aquesta qüestió, s'ha consultat en el *heatmap* de Canvis Percentuals quins són els països que han augmentat aquestes concentracions. D'aquest procés, s'extreuen països de diferents regions: Mongòlia, Tanzània, Kazakhstan, Jordània, Kènia, Armènia, Austràlia, Jamaica entre d'altres.

Per a països com Tanzània i Kènia, un argument que ens podria ajudar a justificar aquest fenomen seria el seu ràpid creixement de la població sense una transició energètica paral·lela; la qual cosa, en aquests països, suposa més crema de

biomassa en termes absoluts³². D'altra banda, a països com Armènia (Caucas del Sud) i Kazakhstan (Àsia Central), les xarxes de gas centralitzades de l'era soviètica es van desmantellar durant la dècada anterior a la nostra finestra temporal. Això, hauria implicat que, al llarg de les darreres dues dècades, la població hagués de tornar a fonts individuals de calefacció, com ara estufes de carbó i llenya.

La relació entre aquest fenomen i els nostres indicadors Matern-Filials es recolzen pel fet que, a Armènia, els costos dels efectes del PM2.5 sobre la salut representen el 10% del GDP d'Armènia³⁵. Encara més, a Kazakhstan, per a un total de 21 ciutats estudiades, s'han registrat fins a 8.134 morts anuals (mitjana de 2015-2017) atribuïdes a PM2.5³⁶.

En aquests casos, però, no trobem factors fonamentals del Canvi Global, sinó desacoblaments entre transicions demogràfiques i energètiques que deriven en aquest augment de les concentracions de PM2.5. Ara bé, també cal esmentar els casos de Mongòlia i Austràlia. Per al primer, sí que es denota una forta industrialització i expansió de la mineria del carbó;³⁷ mentre que per a Austràlia, les xifres de PM2.5 podrien ser causades pels intensos períodes d'incendis que Austràlia ha patit darrerament³⁸.

Així i tot, també hem de tenir en compte l'indicador HidroM. En el *heatmap* de valors mitjans, veiem que pràcticament tots els països amb xifres preocupants respecte a aquest indicador s'ubiquen a l'Àfrica. Encara més, alguns coincideixen amb els països que han augmentat el PM2.5, com són Kènia i Tanzània. El fet que la combinació d'aquests dos indicadors siguin capaços d'explicar un 27,9% de la variància del BPN podria indicar que la combinació de l'obsolescència de la infraestructura energètica i de sanejament de l'aigua podria impactar en la salut de la mare i el nadó.

Per acabar, durant els antecedents hem mencionat la relació entre la radiació UV i la contaminació atmosfèrica. Aquesta relació pren rellevància quan tenim en compte marcs teòrics com la hipòtesi de Jablonski; la qual proposa que l'evolució de la pigmentació cutània s'estableix arran d'un equilibri entre la degradació del folat (Vitamina B9) i la síntesi de vitamina D³⁹. Quant a això, malgrat no trobar cap indici de selecció balancejadora al *PopHuman* sobre el polimorfisme rs1801133, sí que ens trobem que en la distribució de les freqüències al·lèliques s'intueix un patró similar al de la radiació UV. Això podria comportar un altre component sobre el qual ajustar els models estadístics que avaluen la salut reproductiva de la dona en ciutats cosmopolites.

Conclusions

Al llarg d'aquesta anàlisi hem vist com els diferents indicadors del Canvi Global reflecteixen el panorama del desenvolupament en cadascun dels territoris. A més, hem còpsat un model estadístic que explica la variabilitat d'un indicador de la salut materna-filial a partir de dos indicadors prou diferents del Canvi Global. I, finalment, hem validat qualitativament una superposició entre les freqüències al·lèliques i la radiació ultraviolada.

Així i tot, la disponibilitat d'indicadors i la seva qualitat han deixat molt a desitjar. No només ha estat impossible trobar dades fidels de la Salut Reproductiva de la Dona, com ara la prevalença d'endometriosis; sinó que els indicadors de la Salut Materna-Filial tenen una menor representació malgrat aportar, simultàniament, informació de la salut d'un individu i la seva potencial descendència.

En vista d'això, aquest treball també s'orienta amb la intenció d'encoratjar estudis ecològics amb més resolució espacial. La combinació entre evidència epidemiològica d'aquest nivell, juntament amb assajos de laboratori, podria ser una bona manera d'estimular polítiques tant a escala nacional com urbanística; fonamental per a la resiliència de les societats davant els fenòmens del Canvi Climàtic.

Bibliografia

1. Crutzen, Paul J.; Stoermer, Eugene F. The 'Anthropocene'. in *The Future of Nature: Documents of Global Change* 484–486 (Yale University Press, New Haven).
2. Pimentel, David; Burgess, Michael. Soil Erosion Threatens Food Production. **3**, 443–463 (2013).
3. Hawkes, Corinna. Uneven dietary development: linking the policies and processes of globalization with the nutrition transition, obesity and diet-related chronic diseases. **2**, 1–18 (2006).
4. Allen, Myles R. et al. Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature* **458**, 1163–1166 (2009).
5. Ragusa, Antonio et al. Plasticenta: First evidence of microplastics in human placenta. <https://doi.org/10.1016/j.envint.2020.106274> (2021)
doi:10.1016/j.envint.2020.106274.
6. Wild, Christopher Paul. The exposome: from concept to utility. **41**, 24–32 (2012).

7. Barouki, Robert (*), Gluckman, Peter D., Grandjean, Philippe, Hanson, Mark, & Heindel, Jerrold J. Developmental origins of non-communicable disease: Implications for research and public health. **11**, (2012).
8. World Health Organization. Reproductive health: draft strategy to accelerate progress towards the attainment of international development goals and targets. (2004).
9. McCue, Kelly, MD & DeNicola, Nathaniel, MD, MSHP (*). Environmental Exposures in Reproductive Health. **46**, 455–468 (2019).
10. Rumph, Jelonia T. *et al.* Uncovering Evidence: Associations between Environmental Contaminants and Disparities in Women's Health. *International Journal of Environmental Research and Public Health* <https://doi.org/10.3390/ijerph19031257> doi:10.3390/ijerph19031257.
11. Hart, R. J. Physiological Aspects of Female Fertility: Role of the Environment, Modern Lifestyle, and Genetics. *Physiological Reviews*. **96**, 873–909 (2016).
12. Olena Gruzieva, Cheng-Jian Xu & Carrie V. Breton. Epigenome-Wide Meta-Analysis of Methylation in Children Related to Prenatal NO₂ Air Pollution Exposure. **125**, 104–110 (2017).
13. Zhang, Liqiang *et al.* Air pollution-induced missed abortion risk for pregnancies. *Nature Sustainability* 100–109 (2019) doi:10.1038/s41893-019-0387-y.
14. Mortimer, Kathleen, Lurmann, Frederick, Jerrett, Michael, & Tager, Ira B. Exposure to Traffic-related Air Pollution During Pregnancy and Term Low Birth Weight: Estimation of Causal Associations in a Semiparametric Model. *American Journal of Epidemiology* **172**, 815–824 (2012).
15. Stieb, David M., Chen, Li, Eshoul, Maysoon & Judek, Stan. Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. *Environmental Research* **117**, 100–111 (2012).
16. Padula, Amy M. (*) *et al.* Exposure to airborne polycyclic aromatic hydrocarbons during pregnancy and risk of preterm birth. *Environmental Research* **135**, 221–226 (2014).
17. Holstius, David M., Reid, Colleen E., Jesdale, Bill M., & Morello-Frosch, Rachel (*). Weight following Pregnancy during the 2003 Southern California Wildfires. *Environmental Health Perspectives* **120**, 1340–1345 (2012).
18. Heat Exposure and Maternal Health in the Face of Climate Change.

19. He, Siyi, Kosatsky, Tom, Smargiassi, Audrey, Bilodeau-Bertrand, Marianne, & Auger, Nathalie (*). Heat and pregnancy-related emergencies: Risk of placental abruption during hot weather. *Environment International* **111**, 295–300 (2018).
20. Rammah, Amal *et al.* Temperature, placental abruption and stillbirth. *Environment International* **131**, (2019).
21. Chersich, Matthew Francis *et al.* Associations between high temperatures in pregnancy and risk of preterm birth, low birth weight, and stillbirths: systematic review and meta-analysis. *BMJ (British Medical Journal)* **371**, (2020).
22. Ren, C. *et al.* Ambient temperature, air pollution, and heart rate variability in an aging population. *American Journal of Epidemiology*. **173**, 1013–1021 (2011).
23. Jacob, D. J. & Winner, D. A. Effect of climate change on air quality. *Atmospheric Environment*. **43**, 51–63 (2009).
24. Mark Luccock; Emma Beckett; Charlotte Martin; Patrice Jones; John Furst; Zoe Yates; Nina G. Jablonski; George Chaplin; Martin Veysey. UV-associated decline in systemic folate: implications for human nutrigenetics, health, and evolutionary processes. *American Journal of Human Biology* **29**, e22929 (2017).
25. Suarez, Lucina, Felkner, Marilyn, & Hendricks, Kate. The Effect of Fever, Febrile Illnesses, and Heat Exposures on the Risk of Neural Tube Defects in a Texas-Mexico Border Population. *Birth Defects Research (Part A): Clinical and Molecular Teratology* **70**, 815–819 (2004).
26. Canipari, Rita, De Santis, Lucia, & Cecconi, Sandra. Female Fertility and Environmental Pollution . *International Journal of Environmental Research and Public Health* **17**, (2020).
27. Khan, Aneire Ehmar *et al.* Salinity in Drinking Water and the Risk of (Pre)Eclampsia and Gestational Hypertension in Coastal Bangladesh: A Case-Control Study. *PLoS ONE* **9**, (2014).
28. UN WomenWatch / United Nations. *Women, Gender Equality and Climate Change*.
29. Posit team. RStudio: Integrated Development Environment for R. Posit Software, PBC (2025).
30. Wang, Wei; Xie, Xin; Yuan, Ting; Wang, Yanyan; Zhao, Fei; Zhou, Zhangjian; Zhang, Hao. Epidemiological trends of maternal hypertensive disorders of pregnancy at the global, regional, and national levels: a population-based study. **21**, (2021).
31. The jamovi project. jamovi. (2025).

32. International Energy Agency. *Africa Energy Outlook 2022*. <https://www.iea.org/> (2022).
33. World Health Organization. *WHO Ambient Air Quality Database, 2022 Update: Status Report*. (2023).
34. Alikhani Faradonbeh, Mahdijeh; Mardani, Gashtasb; Raeisi Shahraki, Hadi. Longitudinal Trends of the Annual Exposure to PM_{2.5} Particles in European Countries. *Scientifica* **2021**, (2021).
35. World Bank. *Country Climate and Development Report: Armenia*. (2024).
36. Aiyngul Kerimray; Daulet Assanov; Bulat Kenessov; Ferhat Karaca. Trends and health impacts of major urban air pollutants in Kazakhstan. **70**, 1148–1164 (2020).
37. Gheorghe, Adrian; Ankhbayar, Batbayar; van Nieuwenhuyzen, Henlo; de Sa, Rogerio. *Mongolia's Air Pollution Crisis: A Call to Action to Protect Children's Health*. <http://www.unicef.mn/> (2018).
38. Yiwen Zhang; Tingting Ye; Pei Yu; Rongbin Xu; Gongbo Chen; Wenhua Yu; Jiangning Song; Yuming Guo; Shanshan Li. Preterm birth and term low birth weight associated with wildfire-specific PM_{2.5}: A cohort study in New South Wales, Australia during 2016–2019. **174**, 107879 (2023).
39. Nina G. Jablonski; George Chaplin. The evolution of human Skin coloration. **39**, 57–106 (2000).

Annexos

Script 1. Encavalcament de bases de dades primàries

```
1 library(readxl); library(dplyr); library(writexl)
2 # --- 1. SELECCIONAR FITXER D'ENTRADA ---
3 INPUT_FILE <- file.choose()
4 cat("Fitxer seleccionat:", INPUT_FILE, "\n")
5 OUTPUT_FILE <- file.path(getwd(), "SeriesT_Merg.xlsx")
6 cat("Fitxer de sortida:", OUTPUT_FILE, "\n\n")
7 cat("Llegint", INPUT_FILE, "... \n")
8 # --- 2. BASE: Reference (212 països x anys 2000-2020) ---
9 ref <- read_excel(INPUT_FILE, sheet="Reference") %>% rename(Continent=Continente)
10 years <- 2000:2020
11 base <- ref %>% slice(rep(1:n(), each=length(years))) %>% mutate(Year=rep(years, times=nrow(ref)))
12 cat(sprintf(" Base: %d files (%d països x %d anys)\n", nrow(base), nrow(ref), length(years)))
13 ref_codes <- unique(ref$CountryCode)
14 data_sheets <- setdiff(excel_sheets(INPUT_FILE), "Reference")
15 # --- 3. FUSIÓ AUTOMÀTICA ---
16 merged <- base
17 cat("\nFusionant fulles:\n")
18 cat(sprintf("%-14s | països: %3s/%-3s | files: %7s/%-7s | cobertura files: %5s | cobertura països: %5s\n",
19 "Indicador", "ok", "tot", "ok", "tot", "%", "%"))
20 for (sheet in data_sheets) {
21   df_raw <- read_excel(INPUT_FILE, sheet=sheet)
22   cols <- names(df_raw)
23   code_col <- cols[grepl("^CountryCode", cols)][1]
24   year_col <- cols[grepl("^Year", cols)][1]
25   value_col <- cols[grepl("^Value", cols)][1]
26   if (any(is.na(c(code_col, year_col, value_col)))) {
27     cat(sprintf(" AVÍS: fulla '%s' no segueix el patró esperat, omesa.\n", sheet)); next
28   }
29   df <- df_raw %>%
30     select(all_of(c(code_col, year_col, value_col))) %>%
31     rename(CountryCode=all_of(code_col), Year=all_of(year_col), !!sheet:=all_of(value_col)) %>%
32     mutate(Year=as.integer(Year))
33   merged <- merged %>% left_join(df, by=c("CountryCode","Year"))
34   # --- Càlcul i impressió de la cobertura per fulla ---
35   filled <- sum(!is.na(merged[[sheet]]))
36   matched <- length(intersect(ref_codes, unique(df$CountryCode)))
37   coverage_files_pct <- round((filled/nrow(merged))*100, 1)
38   coverage_paisos_pct <- round((matched/length(ref_codes))*100, 1)
39   cat(sprintf(" %-14s | països: %3d/%-3d | files: %5d/%-7d | %7.1f%% | %7.1f%%\n",
40     sheet, matched, length(ref_codes), filled, nrow(merged),
41     coverage_files_pct, coverage_paisos_pct))
42 }
43 # --- 4. DESAR ---
44 write_xlsx(merged, OUTPUT_FILE)
45 cat(sprintf("\nDesat: '%s' (%d files x %d columnes)\n", OUTPUT_FILE, nrow(merged), ncol(merged)))
```

Script 2. Càlcul de Valors Mitjans i coverage

```
1 # =====
2 # Script_valors_medies.R
3 # Calcula la mitjana triennal de cada indicador per país (2000–2020).
4 # P1:2000–2002 | P2:2003–2005 | P3:2006–2008 | P4:2009–2011
5 # P5:2012–2014 | P6:2015–2017 | P7:2018–2020
6 # UV, HDP, HidroM → snapshot | Anomaly, Affecteds → sumatori absoluts
7 # Output: SeriesT_Merg_Average.xlsx
8 # =====
9 library(readxl); library(writexl); library(dplyr); library(tidyr)
10 # --- 1. SELECCIÓ DEL FITXER D'ENTRADA ---
11 fitxer_input <- file.choose()
12 # --- 2. LECTURA I PREPARACIÓ ---
13 dades <- read_excel(fitxer_input)
14 cat("Columnes:", paste(names(dades), collapse=" ", "\n"))
15 cat("Files:", nrow(dades), "| Països:", n_distinct(dades$CountryCode), "\n\n")
16 vars_numeriques <- setdiff(names(dades), c("Continent","CountryName","CountryCode","Year"))
17 dades <- dades %>%
18   mutate(across(all_of(vars_nueriques), ~suppressWarnings(as.numeric(.x)))) %>%
19   mutate(across(all_of(vars_numeriques), ~ifelse(.x==0, NA_real_, .x))) %>%
20   mutate(Periode=paste0("P", ceiling((Year-1999)/3)))
21 # --- 3. CLASSIFICACIÓ DE VARIABLES ---
22 vars_id <- c("Continent","CountryName","CountryCode")
23 vars_snap <- c("UV","HDP","HidroM")
24 vars_sum <- c("Anomaly","Affecteds")
25 vars_avg <- setdiff(names(dades), c(vars_id,"Year","Periode",vars_snap,vars_sum))
26 cat("Mitjana:", paste(vars_avg, collapse=" ", "\n"))
27 cat("Snapshot:", paste(vars_snap, collapse=" ", "\n"))
28 cat("Sumatori:", paste(vars_sum, collapse=" ", "\n\n"))
29 # --- 4. Càlcul de mètriques per país × període ---
30 vars_grup <- c(vars_id,"Periode")
31 df_avg <- dades %>% group_by(across(all_of(vars_grup))) %>%
32   summarise(across(all_of(vars_avg), ~mean(.x, na.rm=TRUE)), .groups="drop")
33 df_snap <- dades %>% arrange(CountryCode, Year) %>% group_by(across(all_of(vars_grup))) %>%
34   summarise(across(all_of(vars_snap), ~{vals<-x[!is.na(.x)]; if(length(vals)==0L) NA_real_ else tail(vals,1L)}),
35   .groups="drop")
36 df_sum <- dades %>% group_by(across(all_of(vars_grup))) %>%
37   summarise(across(all_of(vars_sum), ~{vals<-abs(.x[!is.na(.x)]); if(length(vals)==0L) NA_real_ else sum(vals)}),
38   .groups="drop")
39 ordre_final <- c(vars_id,"Periode",vars_avg,vars_snap,vars_sum)
40 resultat_periodes <- df_avg %>%
41   left_join(df_snap, by=vars_grup) %>% left_join(df_sum, by=vars_grup) %>%
42   select(all_of(ordre_final))
43 # --- 4e. Mitjana global (TOTAL) ---
44 total_avg <- dades %>% group_by(across(all_of(vars_id))) %>%
45   summarise(across(all_of(vars_avg), ~mean(.x, na.rm=TRUE)), .groups="drop")
46 total_snap <- dades %>% arrange(CountryCode, Year) %>% group_by(across(all_of(vars_id))) %>%
47   summarise(across(all_of(vars_snap), ~{vals<-x[!is.na(.x)]; if(length(vals)==0L) NA_real_ else tail(vals,1L)}),
48   .groups="drop")
49 total_sum <- dades %>% group_by(across(all_of(vars_id))) %>%
50   summarise(across(all_of(vars_sum), ~{vals<-abs(.x[!is.na(.x)]); if(length(vals)==0L) NA_real_ else sum(vals)}),
51   .groups="drop")
52 resultat_total <- total_avg %>%
53   left_join(total_snap, by=vars_id) %>% left_join(total_sum, by=vars_id) %>%
54   mutate(Periode="TOTAL") %>% select(all_of(ordre_final))
55 resultat <- bind_rows(resultat_total, resultat_periodes) %>%
56   arrange(CountryCode, match(Periode, c("TOTAL", paste0("P",1:7)))) %>%
57   mutate(across(where(is.numeric), ~ifelse(is.nan(.x), NA_real_, .x)))
58 cat("Resultat:", nrow(resultat), "files |", n_distinct(resultat$CountryCode),
59   "països ×", n_distinct(resultat$Periode), "períodes\n\n")
60 # --- 5. TAULA DE COBERTURA ---
61 tots_indicadors <- c(vars_avg, vars_snap, vars_sum)
62 n_total_pais <- n_distinct(dades$CountryCode)
63 periodes_ord <- paste0("P",1:7)
64 cob_global <- laply(tots_indicadors, function(v) {
65   resum <- dades %>% group_by(CountryCode) %>%
66     summarise(files_dades=sum(!is.na(.data[[v]])), .groups="drop")
67   data.frame(Indicador=v,
68     Tipus=ifelse(v%in%vars_snap,"Snapshot",ifelse(v%in%vars_sum,"Sumatori","Mitjana")),
```

```

64   Periode="TOTAL", Files_amb_dades=sum(resum$files_dades), Total_files=nrow(dades),
65   Cobertura_files_pct=round(100*sum(resum$files_dades)/nrow(dades),1),
66   Paisos_amb_dades=sum(resum$files_dades>0), Total_paisos=n_total_pais,
67   Cobertura_pais_pct=round(100*sum(resum$files_dades>0)/n_total_pais,1), stringsAsFactors=FALSE)
68   }) %>% bind_rows()
69   cob_periode <- lapply(tots_indicadors, function(v) {
70     lapply(periodes_ord, function(p) {
71       sub <- dades %>% filter(Periode==p)
72       resum <- sub %>% group_by(CountryCode) %>%
73         summarise(files_dades=sum(!is.na(.data[[v]])), .groups="drop")
74       data.frame(Indicador=v,
Tipus=ifelse(v%in%vars_snap,"Snapshot",ifelse(v%in%vars_sum,"Sumatori","Mitjana")),
75       Periode=p, Files_amb_dades=sum(resum$files_dades), Total_files=nrow(sub),
76       Cobertura_files_pct=round(100*sum(resum$files_dades)/nrow(sub),1),
77       Paisos_amb_dades=sum(resum$files_dades>0), Total_paisos=n_total_pais,
78       Cobertura_pais_pct=round(100*sum(resum$files_dades>0)/n_total_pais,1), stringsAsFactors=FALSE)
79     }) %>% bind_rows()
80   }) %>% bind_rows()
81   cobertura <- bind_rows(cob_global, cob_periode) %>%
82     arrange(Indicador, match(Periode, c("TOTAL", periodes_ord)))
83   # --- 6. EXPORTACIÓ ---
84   fitxer_output <- file.path(getwd(), "SeriesT_Merg_Average.xlsx")
85   writexl::write_xlsx(resultat, path=fitxer_output)
86   cat("=====\n")
87   cat("Fitxer guardat a:\n", fitxer_output, "\n")
88   cat(" ", nrow(resultat), "files |", n_distinct(resultat$CountryCode),
89     "països |", n_distinct(resultat$Periode), "períodes\n")
90   cat("=====\n")
91   # --- 7. TAULA CREUADA: % cobertura per indicador × període ---
92   cat("\n--- Taula creuada: % cobertura per indicador × període ---\n")
93   tab_ind_periode <- cob_periode %>%
94     select(Indicador, Periode, Cobertura_pais_pct) %>%
95     pivot_wider(names_from=Periode, values_from=Cobertura_pais_pct) %>%
96     select(Indicador, all_of(periodes_ord)) %>%
97     left_join(cob_global %>% select(Indicador, Països=Cobertura_pais_pct, Files=Cobertura_files_pct),
by="Indicador") %>%
98     select(Indicador, all_of(periodes_ord), Països, Files)
99     print(as.data.frame(tab_ind_periode), row.names=FALSE)

```

Script 3. Càlcul de Canvis Percentuals i coverage

```
1      # =====
2      # Script: pct_change_series_temporals.R
3      # Calcula el canvi percentual entre el primer i últim any amb
4      # dada disponible per a cada país i variable. Detecta automàti-
5      # cament indicadors snapshot i els inclou com a valor discret.
6      # Anomaly, Affecteds: suma de valors absoluts per país.
7      # Input: SeriesT_Merg.xlsx | Output: SeriesT_Merg_PctC.xlsx
8      # =====
9      library(readxl); library(dplyr); library(writexl)
10     # --- 1. PARÀMETRES ---
11     input_file <- "Series_Temporals_Merged.xlsx"
12     output_file <- "Series_Temporals_PctChange.xlsx"
13     # --- 2. CARREGAR DADES ---
14     merged <- read_excel(input_file)
15     cat("Dades carregades:", nrow(merged), "files,", ncol(merged), "columnes\n")
16     id_cols <- c("Continent","CountryName","CountryCode","Year")
17     var_cols <- setdiff(names(merged), id_cols)
18     cat("Variables detectades:", paste(var_cols, collapse=" "), "\n\n")
19     # --- 3. CASOS ESPECIALS ---
20     abssum_vars <- c("Anomaly","Affecteds")
21     var_cols_general <- setdiff(var_cols, abssum_vars)
22     # --- 4. DETECTAR INDICADORS SNAPSHOT ---
23     is_snapshot <- function(var) {
24       df_var <- merged %>% filter(!is.na(.data[[var]])) %>%
25         group_by(CountryCode) %>% summarise(n_years=n_distinct(Year), .groups="drop")
26       if (any(df_var$n_years>1)) return(FALSE)
27       length(unique(merged %>% filter(!is.na(.data[[var]])) %>% pull(Year)))==1
28     }
29     snapshot_vars <- var_cols_general[sapply(var_cols_general, is_snapshot)]
30     temporal_vars <- setdiff(var_cols_general, snapshot_vars)
31     cat("Temporals (PctChange):", paste(temporal_vars, collapse=" "), "\n")
32     cat("Snapshot:", if(length(snapshot_vars)>0) paste(snapshot_vars, collapse=" ") else "(cap)", "\n")
33     cat("Suma absoluta:", paste(abssum_vars, collapse=" "), "\n\n")
34     # --- 5. CALCULAR CANVI PERCENTUAL ---
35     # Ordenat per Year per garantir ordre cronològic correcte
36     pct_change <- function(x, years) {
37       valid <- !is.na(x)
38       if (sum(valid)<2) return(NA_real_)
39       vals <- x[valid][order(years[valid])]
40       if (vals[1]==0) return(NA_real_)
41       (vals[length(vals)]-vals[1])/abs(vals[1])*100
42     }
43     result <- merged %>% group_by(Continent, CountryName, CountryCode) %>%
44       summarise(across(all_of(temporal_vars), ~pct_change(x, Year), .names="{.col}_PctChange"),
45     groups="drop")
46     # --- 6. AFEGIR VALORS SNAPSHOT ---
47     if (length(snapshot_vars)>0) {
48       result_snapshot <- merged %>% group_by(Continent, CountryName, CountryCode) %>%
49         summarise(across(all_of(snapshot_vars), ~na.omit(.x)[1]), .groups="drop")
50       result <- result %>% left_join(result_snapshot, by=c("Continent","CountryName","CountryCode"))
51     }
52     # --- 7. AFEGIR SUMA DE VALORS ABSOLUTS ---
53     result_abssum <- merged %>% group_by(Continent, CountryName, CountryCode) %>%
54       summarise(across(all_of(abssum_vars),
55         ~if(all(is.na(.x))) NA_real_ else sum(abs(.x), na.rm=TRUE), .names="{.col}_AbsSum", .groups="drop")
56     result <- result %>%
57       left_join(result_abssum, by=c("Continent","CountryName","CountryCode")) %>%
58       arrange(Continent, CountryName)
59     cat("Càlcul completat:", nrow(result), "països\n")
60     # --- 8. RESUM DE COBERTURA ---
61     cat("\n--- Cobertura per variable ---\n")
62     for (var in temporal_vars) {
63       col <- paste0(var,"_PctChange"); n <- sum(!is.na(result[[col]]))
64       cat(sprintf(" %-30s: %4d països (%.1f%%)\n", col, n, 100*n/nrow(result)))
65     }
66     for (var in snapshot_vars) {
67       n <- sum(!is.na(result[[var]]))
68       cat(sprintf(" %-30s: %4d països (%.1f%%) [snapshot]\n", var, n, 100*n/nrow(result)))
69     }
```

```
68 }
69 for (var in abssum_vars) {
70   col <- paste0(var, "_AbsSum"); n <- sum(!is.na(result[[col]]))
71   cat(sprintf("%-30s: %4d países (%.1f%%) [suma absoluta]\n", col, n, 100*n/nrow(result)))
72 }
73 # --- 9. EXPORTAR ---
74 write_xlsx(result, output_file)
75 cat("\nFitxer guardat a:", output_file, "\n")
```

Script 4. Heatmap Valors Mitjans

```
1      # Heatmap: Països x Variables – Valors mitjans (període TOTAL, escalat Z-score)
2      # Paquets: readxl, pheatmap, RColorBrewer
3      library(readxl); library(pheatmap); library(RColorBrewer)
4      # --- 1. CÀRREGA DE DADES ---
5      df_all <- as.data.frame(read_excel(file.choose(), sheet=1))
6      df <- df_all[df_all$Període == "TOTAL", ]
7      # --- 2. PREPARACIÓ DE LA MATRIU ---
8      var_cols <- c("CO2","PM2.5","Occupation","Pesticides","HDI","Poisoning",
9                  "BMI","Stillbirth","LBW","Temp","UV","HDP","HidroM","Anomaly","Affecteds")
10     mat <- apply(df[, var_cols], 2, as.numeric)
11     rownames(mat) <- df$CountryName
12     colnames(mat) <- c("CO2_Average","PM2.5_Average","Occupation_Average","Pesticides_Average",
13                     "HDI_Average","Poisoning_Average","BMI_Average","Stillbirth_Average",
14                     "LBW_Average","Temp_AbsSum","UV_Snapshot","HDP_Snapshot",
15                     "HidroM_Snapshot","Anomaly_AbsSum","Affecteds_AbsSum")
16     mat <- mat[rowSums(!is.na(mat)) > 0, ]
17     # --- Imputació de NAs amb la mitjana de columna ---
18     for (j in seq_len(ncol(mat))) {
19       na_idx <- is.na(mat[, j])
20       if (any(na_idx)) mat[na_idx, j] <- mean(mat[, j], na.rm=TRUE)
21     }
22     # --- 3. ESCALAT Z-SCORE PER COLUMNA ---
23     mat_scaled <- scale(mat)
24     # --- 4. ANOTACIÓ DE FILES (continent) ---
25     continent_df <- data.frame(
26       Regió = df$Continent[match(rownames(mat_scaled), df$CountryName)],
27       row.names = rownames(mat_scaled))
28     n_cont <- length(unique(continent_df$Regió))
29     continent_colors <- list(Regió=setNames(
30       brewer.pal(max(3, n_cont), "Set2")[seq_len(n_cont)],
31       sort(unique(continent_df$Regió))))
32     # --- 5. HEATMAP ---
33     pheatmap(
34       mat_scaled,
35       color      = colorRampPalette(c("#2166AC","white","#D6604D"))(100),
36       clustering_method = "ward.D2",
37       cluster_rows=TRUE, cluster_cols=TRUE,
38       annotation_row=continent_df, annotation_colors=continent_colors,
39       show_rownames=TRUE, show_colnames=TRUE,
40       fontsize_row=5, fontsize_col=9, angle_col=45, border_color=NA,
41       main      = "Heatmap dels Valors Mitjans (Escalat Z-score)",
42       filename = file.path(getwd(), "Heatmap_Average.pdf"),
43       width=14, height=22)
44     message("Desat: Heatmap_Average.pdf")
```

Script 5. Heatmap Canvis Percentuals

```
1      # Heatmap: Països x Variables – Canvis percentuals (escalat Z-score)
2      # Paquets: readxl, pheatmap, RColorBrewer
3      library(readxl); library(pheatmap); library(RColorBrewer)
4      # --- 1. CÀRREGA DE DADES ---
5      df <- as.data.frame(read_excel(file.choose(), sheet=1))
6      # --- 2. PREPARACIÓ DE LA MATRIU ---
7      var_cols <- c("CO2_PctChange","PM2.5_PctChange","Occupation_PctChange",
8                  "Pesticides_PctChange","HDI_PctChange","BMI_PctChange",
9                  "Stillbirth_PctChange","LBW_PctChange","Poisoning_PctChange",
10                 "UV","HDP","HidroM","Anomaly_AbsSum","Affecteds_AbsSum","Temp_AbsSum")
11      mat <- apply(df[, var_cols], 2, as.numeric)
12      rownames(mat) <- df$CountryName
13      colnames(mat)[colnames(mat)=="UV"] <- "UV_Snapshot"
14      colnames(mat)[colnames(mat)=="HDP"] <- "HDP_Snapshot"
15      colnames(mat)[colnames(mat)=="HidroM"] <- "HidroM_Snapshot"
16      mat <- mat[rowSums(!is.na(mat)) > 0, ]
17      # --- Imputació de NAs amb la mitjana de columna ---
18      for (j in seq_len(ncol(mat))) {
19          na_idx <- is.na(mat[, j])
20          if (any(na_idx)) mat[na_idx, j] <- mean(mat[, j], na.rm=TRUE)
21      }
22      # --- 3. ESCALAT Z-SCORE PER COLUMNA ---
23      mat_scaled <- scale(mat)
24      # --- 4. ANOTACIÓ DE FILES (continent) ---
25      continent_df <- data.frame(
26          Regió = df$Continent[match(rownames(mat_scaled), df$CountryName)],
27          row.names = rownames(mat_scaled))
28      n_cont <- length(unique(continent_df$Regió))
29      continent_colors <- list(Regió=setNames(
30          brewer.pal(max(3, n_cont), "Set2")[seq_len(n_cont)],
31          sort(unique(continent_df$Regió))))
32      # --- 5. HEATMAP ---
33      pheatmap(
34          mat_scaled,
35          color = colorRampPalette(c("#2166AC","white","#D6604D"))(100),
36          clustering_method = "ward.D2",
37          cluster_rows=TRUE, cluster_cols=TRUE,
38          annotation_row=continent_df, annotation_colors=continent_colors,
39          show_rownames=TRUE, show_colnames=TRUE,
40          fontsize_row=5, fontsize_col=9, angle_col=45, border_color=NA,
41          main = "Heatmap dels Canvis Percentuals (Escalat Z-score)",
42          filename = file.path(getwd(), "Heatmap_PctChange.pdf"),
43          width=14, height=22)
44      message("Desat: Heatmap_PctChange.pdf")
```

Script 6. Bi-plot i Scree-plot Valors Mitjans

```
1 # PCA – Dades de mitjanes: Screeplot + Biplots 2D (PC1–PC3)
2 # Scores pre-calculats pel Jamovi. Variables suplementàries projectades per correlació.
3 # Paquets: readxl, ggplot2, ggrepel, RColorBrewer, cluster
4 library(readxl); library(ggplot2); library(ggrepel); library(RColorBrewer); library(cluster)
5 # --- 1. DADES ---
6 df <- as.data.frame(read_excel(file.choose(), sheet=1))
7 pc_cols <- c("PC1","PC2","PC3")
8 sup_cols <- c("Stillbirth","LBW","HDP","HDI")
9 # --- 2. COLORS ---
10 region_labels <- c(AFR="AFR – Àfrica", AMR="AMR – Amèriques", EMR="EMR – Mediterrani Oriental",
11   EUR="EUR – Europa", SEAR="SEAR – Àsia Sud-Est", WPR="WPR – Pacífic Occidental")
12 region_colors <- setNames(brewer.pal(6,"Set2"), unname(region_labels))
13 df$Region <- unname(region_labels[df$Continent])
14 # --- 3. SCREEPLOT ---
15 eig_df <- as.data.frame(read_excel(file.choose(), sheet=2, skip=1))
16 pct_var_all <- as.numeric(eig_df[[3]])
17 cum_var_all <- as.numeric(eig_df[[4]])
18 all_eig <- as.numeric(eig_df[[2]]); all_eig[is.na(all_eig)]
19 n_show <- 5; pct_var <- pct_var_all
20 kaiser_pct <- 100 / length(all_eig)
21 scree_df <- data.frame(
22   PC = factor(paste0("PC",1:n_show), levels=paste0("PC",1:n_show)),
23   Variància = pct_var_all[1:n_show],
24   Acumulada = cum_var_all[1:n_show])
25 p_scree <- ggplot(scree_df, aes(x=PC)) +
26   geom_col(aes(y=Variància), fill="#8DA0CB", alpha=0.85, width=0.6) +
27   geom_line(aes(y=Acumulada, group=1), color="#E41A1C", linewidth=0.8) +
28   geom_point(aes(y=Acumulada), color="#E41A1C", size=2.5) +
29   geom_hline(yintercept=kaiser_pct, linetype="dashed", color="#2CA02C", linewidth=0.8) +
30   annotate("text", x=n_show+0.45, y=kaiser_pct+2.5,
31     label="Criteri de Kaiser (eigenvalor = 1)", hjust=1, vjust=0, size=3, color="#2CA02C") +
32   scale_y_continuous(name="Variància explicada per component (%)",
33     sec.axis=sec_axis(~., name="Variància acumulada (%)"),
34     limits=c(0,100), expand=expansion(mult=c(0,0.02)),
35     labels=function(x) format(x, decimal.mark="")) +
36   labs(title="Screeplot – PCA de Valors Mitjans", x="Component Principal",
37     caption="Barres blaves: variància per component. Punts i línia vermells: variància acumulada. Línia verda:
38   criteri de Kaiser.") +
39   theme_minimal(base_size=12) +
40   theme(plot.title=element_text(face="bold", size=14, margin=margin(b=24)),
41     plot.caption=element_text(size=8, color="grey50"),
42     plot.margin=margin(1,1.5,1,1,"cm"),
43     axis.title.x=element_text(margin=margin(t=14)),
44     axis.title.y=element_text(margin=margin(r=14)),
45     axis.title.y.right=element_text(margin=margin(l=14)),
46     axis.line=element_line(color="black", linewidth=0.5),
47     panel.grid.major=element_line(color="grey85", linewidth=0.4),
48     panel.grid.minor=element_line(color="grey92", linewidth=0.25))
49 ggsave(file.path(getwd(),"Screeplot_PCA_Average.pdf"), p_scree, width=7, height=5)
50 message("Desat: Screeplot_PCA_Average.pdf")
51 # --- 4. VARIABLES SUPLEMENTÀRIES: PROJECCIÓ PER CORRELACIÓ ---
52 scores_mat_all <- as.matrix(df[, pc_cols])
53 sup_mat <- as.matrix(df[, sup_cols])
54 complete_idx <- complete.cases(scores_mat_all)
55 cor_mat_all <- cor(sup_mat[complete_idx,], scores_mat_all[complete_idx,], use="pairwise.complete.obs")
56 # --- 5. BIPLOTS 2D PER A TOTES LES COMBINACIONS DE PC1–PC3 ---
57 scores_df <- df[complete_idx,]
58 pc_pairs <- combn(pc_cols, 2, simplify=FALSE)
59 pdf(file.path(getwd(),"Biplots_PCA_Average.pdf"), width=10, height=8)
60 for(pair in pc_pairs){
61   px <- pair[1]; py <- pair[2]
62   px_idx <- which(pc_cols==px); py_idx <- which(pc_cols==py)
63   sc_mat_2d <- as.matrix(scores_df[, pair])
64   cor_pair <- cor_mat_all[, c(px_idx,py_idx)]
65   scale_f <- max(abs(sc_mat_2d), na.rm=TRUE) / max(abs(cor_pair), na.rm=TRUE) * 0.75
66   load_pair <- as.data.frame(cor_pair * scale_f)
67   colnames(load_pair) <- c("Xvar","Yvar"); load_pair$Variable <- sup_cols
68   # --- K-means amb k òptim per silhouette ---
```

```

68 sil_vals <- sapply(2:6, function(k) {
69   km <- kmeans(sc_mat_2d, centers=k, nstart=25, iter.max=100)
70   mean(silhouette(km$cluster, dist(sc_mat_2d))[,3]))
71   optimal_k <- which.max(sil_vals) + 1
72   km_fit <- kmeans(sc_mat_2d, centers=optimal_k, nstart=25, iter.max=100)
73   scores_df_p <- scores_df
74   scores_df_p$Cluster <- factor(paste0("C", km_fit$cluster))
75   scores_df_p$Xvar <- scores_df[[px]]; scores_df_p$Yvar <- scores_df[[py]]
76   centroids_df <- data.frame(Xvar=km_fit$centers[,1], Yvar=km_fit$centers[,2],
77     Cluster=paste0("C", seq_len(optimal_k)))
78   px_pct <- round(pct_var[px_idx], 1); py_pct <- round(pct_var[py_idx], 1)
79   p_bip <- ggplot() +
80     stat_ellipse(data=scores_df_p, aes(x=Xvar, y=Yvar, group=Cluster),
81       geom="polygon", level=0.95, alpha=0.08, color="grey35", fill="grey70",
82       linetype="dashed", linewidth=0.5) +
83     geom_label(data=centroids_df, aes(x=Xvar, y=Yvar, label=Cluster),
84       size=3.5, color="grey20", fill="white", alpha=0.75, label.size=0.3) +
85     geom_point(data=scores_df_p, aes(x=Xvar, y=Yvar, color=Region), size=2, alpha=0.8) +
86     geom_segment(data=load_pair, aes(x=0, y=0, xend=Xvar, yend=Yvar),
87       arrow=arrow(length=unit(0.25,"cm")), type="closed", color="grey20", linewidth=0.5) +
88     geom_text_repel(data=load_pair, aes(x=Xvar, y=Yvar, label=Variable),
89       size=3, color="grey20", max.overlaps=20) +
90     scale_color_manual(values=region_colors) +
91     geom_hline(yintercept=0, linetype="dashed", color="grey60", linewidth=0.4) +
92     geom_vline(xintercept=0, linetype="dashed", color="grey60", linewidth=0.4) +
93     labs(title=paste0("Biplot PCA – Valors Mitjans (" , px, " vs ", py, ")"),
94       subtitle=paste0(px, ": ", px_pct, "% | ", py, ": ", py_pct, "% | Total: ", round(px_pct+py_pct, 1), "%"),
95       x=paste0(px, " (" , px_pct, "%)"), y=paste0(py, " (" , py_pct, "%)"), color="Regió",
96       caption="Font: PCA-PERMANOVA_Average_Data.xlsx. Fletxes: variables suplementàries projectades per
correlació.") +
97     theme_minimal(base_size=12) +
98     theme(plot.title=element_text(face="bold", size=14, margin=margin(b=8)),
99       plot.subtitle=element_text(size=10, color="grey40", margin=margin(b=16)),
100       plot.caption=element_text(size=8, color="grey50"),
101       plot.margin=margin(1, 1, 1, 1, "cm"),
102       axis.title.x=element_text(margin=margin(t=14)),
103       axis.title.y=element_text(margin=margin(r=14)),
104       axis.line=element_line(color="black", linewidth=0.5),
105       panel.grid.major=element_line(color="grey85", linewidth=0.4),
106       panel.grid.minor=element_line(color="grey92", linewidth=0.25),
107       legend.position="bottom")
108   print(p_bip)
109   message("Biplot: ", px, " vs ", py, " (k = ", optimal_k, ")")
110 }
111 dev.off()
112 message("Desat: Biplots_PCA_Average.pdf")

```

Script 7. Bi-plot i Scree-plots Canvis Percentuals

```
1      # PCA – Canvis Percentuals: Screeplot + Biplots 2D (Dim1–Dim4)
2      # Scores pre-calculats pel Jamovi. Variables suplementàries projectades per correlació.
3      # Paquets: readxl, ggplot2, ggrepel, RColorBrewer, cluster
4      library(readxl); library(ggplot2); library(ggrepel); library(RColorBrewer); library(cluster)
5      # --- 1. DADES ---
6      df <- as.data.frame(read_excel(file.choose(), sheet=1))
7      dim_cols <- c("Dim1","Dim2","Dim3","Dim4")
8      sup_cols <- c("Stillbirth_PctChange","LBW_PctChange","HDP","HDI_PctChange")
9      # --- 2. COLORS ---
10     region_labels <- c(AFR="AFR – Àfrica", AMR="AMR – Amèriques", EMR="EMR – Mediterrani Oriental",
11                        EUR="EUR – Europa", SEAR="SEAR – Àsia Sud-Est", WPR="WPR – Pacífic Occidental")
12     region_colors <- setNames(brewer.pal(6,"Set2"), unname(region_labels))
13     df$Region <- unname(region_labels[df$Continent])
14     # --- 3. SCREEPLOT ---
15     eig_df <- as.data.frame(read_excel(file.choose(), sheet=2, skip=1))
16     pct_var_all <- as.numeric(eig_df[[3]])
17     cum_var_all <- as.numeric(eig_df[[4]])
18     all_eig <- as.numeric(eig_df[[2]]); all_eig[is.na(all_eig)]
19     n_show <- 5; pct_var <- pct_var_all
20     kaiser_pct <- 100 / length(all_eig)
21     scree_df <- data.frame(
22       PC = factor(paste0("Dim",1:n_show), levels=paste0("Dim",1:n_show)),
23       Variància = pct_var_all[1:n_show],
24       Acumulada = cum_var_all[1:n_show])
25     p_scree <- ggplot(scree_df, aes(x=PC)) +
26       geom_col(aes(y=Variància), fill="#FC8D62", alpha=0.85, width=0.6) +
27       geom_line(aes(y=Acumulada, group=1), color="#E41A1C", linewidth=0.8) +
28       geom_point(aes(y=Acumulada), color="#E41A1C", size=2.5) +
29       geom_hline(yintercept=kaiser_pct, linetype="dashed", color="#2CA02C", linewidth=0.8) +
30       annotate("text", x=n_show+0.45, y=kaiser_pct+2.5,
31              label="Criteri de Kaiser (eigenvalor = 1)", hjust=1, vjust=0, size=3, color="#2CA02C") +
32       scale_y_continuous(name="Variància explicada per component (%)",
33                          sec.axis=sec_axis(~., name="Variància acumulada (%)"),
34                          limits=c(0,100), expand=expansion(mult=c(0,0.02)),
35                          labels=function(x) format(x, decimal.mark="")) +
36       labs(title="Screeplot – PCA de Canvis Percentuals", x="Component Principal",
37            caption="Barres taronges: variància per component. Punts i línia vermells: variància acumulada. Línia verda: criteri de Kaiser.") +
38       theme_minimal(base_size=12) +
39       theme(plot.title=element_text(face="bold", size=14, margin=margin(b=24)),
40            plot.caption=element_text(size=8, color="grey50"),
41            plot.margin=margin(1,1.5,1,1,"cm"),
42            axis.title.x=element_text(margin=margin(t=14)),
43            axis.title.y=element_text(margin=margin(r=14)),
44            axis.title.y.right=element_text(margin=margin(l=14)),
45            axis.line=element_line(color="black", linewidth=0.5),
46            panel.grid.major=element_line(color="grey85", linewidth=0.4),
47            panel.grid.minor=element_line(color="grey92", linewidth=0.25))
48     ggsave(file.path(getwd()),"Screeplot_PCA_PctC.pdf", p_scree, width=7, height=5)
49     message("Desat: Screeplot_PCA_PctC.pdf")
50     # --- 4. VARIABLES SUPLEMENTÀRIES: PROJECCIÓ PER CORRELACIÓ ---
51     scores_mat_all <- as.matrix(df[, dim_cols])
52     sup_mat <- as.matrix(df[, sup_cols])
53     complete_idx <- complete.cases(scores_mat_all)
54     cor_mat_all <- cor(sup_mat[complete_idx,], scores_mat_all[complete_idx,], use="pairwise.complete.obs")
55     # --- 5. BIPLOTS 2D PER A TOTES LES COMBINACIONS DE Dim1–Dim4 ---
56     scores_df <- df[complete_idx,]
57     dim_pairs <- combn(dim_cols, 2, simplify=FALSE)
58     pdf(file.path(getwd()),"Biplots_PCA_PctC.pdf", width=10, height=8)
59     for (pair in dim_pairs) {
60       px <- pair[1]; py <- pair[2]
61       px_idx <- which(dim_cols==px); py_idx <- which(dim_cols==py)
62       sc_mat_2d <- as.matrix(scores_df[, pair])
63       cor_pair <- cor_mat_all[, c(px_idx,py_idx)]
64       scale_f <- max(abs(sc_mat_2d), na.rm=TRUE) / max(abs(cor_pair), na.rm=TRUE) * 0.75
65       load_pair <- as.data.frame(cor_pair * scale_f)
66       colnames(load_pair) <- c("Xvar","Yvar"); load_pair$Variable <- sup_cols
67       # --- K-means amb k òptim per silhouette ---
```

```

68 sil_vals <- sapply(2:6, function(k) {
69   km <- kmeans(sc_mat_2d, centers=k, nstart=25, iter.max=100)
70   mean(silhouette(km$cluster, dist(sc_mat_2d))[,3]))
71   optimal_k <- which.max(sil_vals) + 1
72   km_fit <- kmeans(sc_mat_2d, centers=optimal_k, nstart=25, iter.max=100)
73   scores_df_p <- scores_df
74   scores_df_p$Cluster <- factor(paste0("C", km_fit$cluster))
75   scores_df_p$Xvar <- scores_df[[px]]; scores_df_p$Yvar <- scores_df[[py]]
76   centroids_df <- data.frame(Xvar=km_fit$centers[,1], Yvar=km_fit$centers[,2],
77     Cluster=paste0("C", seq_len(optimal_k)))
78   px_pct <- round(pct_var[px_idx], 1); py_pct <- round(pct_var[py_idx], 1)
79   p_bip <- ggplot() +
80     stat_ellipse(data=scores_df_p, aes(x=Xvar, y=Yvar, group=Cluster),
81       geom="polygon", level=0.95, alpha=0.08, color="grey35", fill="grey70",
82       linetype="dashed", linewidth=0.5) +
83     geom_label(data=centroids_df, aes(x=Xvar, y=Yvar, label=Cluster),
84       size=3.5, color="grey20", fill="white", alpha=0.75, label.size=0.3) +
85     geom_point(data=scores_df_p, aes(x=Xvar, y=Yvar, color=Region), size=2, alpha=0.8) +
86     geom_segment(data=load_pair, aes(x=0, y=0, xend=Xvar, yend=Yvar),
87       arrow=arrow(length=unit(0.25,"cm")), type="closed", color="grey20", linewidth=0.5) +
88     geom_text_repel(data=load_pair, aes(x=Xvar, y=Yvar, label=Variable),
89       size=3, color="grey20", max.overlaps=20) +
90     scale_color_manual(values=region_colors) +
91     geom_hline(yintercept=0, linetype="dashed", color="grey60", linewidth=0.4) +
92     geom_vline(xintercept=0, linetype="dashed", color="grey60", linewidth=0.4) +
93     labs(title=paste0("Biplot PCA – Canvis Percentuals (" ,px," vs " ,py,")"),
94       subtitle=paste0(px," : " ,px_pct,"% | " ,py," : " ,py_pct,"% | Total: " ,round(px_pct+py_pct,1),"%"),
95       x=paste0(px," (" ,px_pct,"%)"), y=paste0(py," (" ,py_pct,"%)"), color="Regió",
96       caption="Font: PCA-PERMANOVA_PctC_Data.xlsx. Fletxes: variables suplementàries projectades per
correlació.") +
97     theme_minimal(base_size=12) +
98     theme(plot.title=element_text(face="bold", size=14, margin=margin(b=8)),
99       plot.subtitle=element_text(size=10, color="grey40", margin=margin(b=16)),
100       plot.caption=element_text(size=8, color="grey50"),
101       plot.margin=margin(1,1,1,1,"cm"),
102       axis.title.x=element_text(margin=margin(t=14)),
103       axis.title.y=element_text(margin=margin(r=14)),
104       axis.line=element_line(color="black", linewidth=0.5),
105       panel.grid.major=element_line(color="grey85", linewidth=0.4),
106       panel.grid.minor=element_line(color="grey92", linewidth=0.25),
107       legend.position="bottom")
108   print(p_bip)
109   message("Biplot: " , px, " vs " , py, " (k = " , optimal_k, ")")
110 }
111 dev.off()
112 message("Desat: Biplots_PCA_PctC.pdf")

```

Script 8. PERMANOVA

```
1 library(vegan); library(dplyr); library(readxl)
2 # --- 1. CÀRREGA DE DADES ---
3 INPUT_FILE <- file.choose()
4 df <- read_excel(INPUT_FILE)
5 vars_totes <- c("Stillbirth_PctChange","LBW_PctChange","HDP",
6 "Dim1","Dim2","Dim3","Dim4","HDI_PctChange","Continent")
7 df_final <- df[complete.cases(df[, vars_totes]), ]
8 cat("Files originals:", nrow(df), "\n")
9 cat("Files completes:", nrow(df_final), "\n\n")
10 paisos_regio <- df_final %>%
11   group_by(Continent) %>% summarise(n_paisos=n(), .groups="drop") %>% arrange(Continent)
12 # --- 2. MATRIU D'OUTCOMES ---
13 Y_repro <- as.matrix(df_final[, c("Stillbirth_PctChange","LBW_PctChange","HDP")])
14 # --- 3. MODELS PERMANOVA (adonis2, by="margin", 999 permutacions) ---
15 set.seed(123)
16 model1 <- adonis2(Y_repro ~ Dim1+Dim2+Dim3+Dim4,
17   data=df_final, permutations=999, method="euclidean", by="margin")
18 set.seed(123)
19 model2 <- adonis2(Y_repro ~ Dim1+Dim2+Dim3+Dim4+HDI_PctChange,
20   data=df_final, permutations=999, method="euclidean", by="margin")
21 set.seed(123)
22 model3 <- adonis2(Y_repro ~ Dim1+Dim2+Dim3+Dim4+Continent,
23   data=df_final, permutations=999, method="euclidean", by="margin")
24 set.seed(123)
25 model4 <- adonis2(Y_repro ~ Dim1+Dim2+Dim3+Dim4+HDI_PctChange+Continent,
26   data=df_final, permutations=999, method="euclidean", by="margin")
27 # --- 4. RESULTATS PER MODEL ---
28 cat("\n===== \n")
29 cat("MODEL 1: Repro ~ Dim1+Dim2+Dim3+Dim4\n")
30 cat("===== \n"); print(model1)
31 cat("\n===== \n")
32 cat("MODEL 2: Repro ~ Dim1+Dim2+Dim3+Dim4+HDI\n")
33 cat("===== \n"); print(model2)
34 cat("\n===== \n")
35 cat("MODEL 3: Repro ~ Dim1+Dim2+Dim3+Dim4+Continent\n")
36 cat("===== \n"); print(model3)
37 cat("\n===== \n")
38 cat("MODEL 4: Repro ~ Dim1+Dim2+Dim3+Dim4+HDI+Continent\n")
39 cat("===== \n"); print(model4)
40 # --- 5. R² AJUSTADA (via capscale) ---
41 cat("\n===== \n")
42 cat("R² ajustada dels models\n")
43 cat("===== \n")
44 cat(sprintf("Model 1 adj.R2: %.4f\n", RsquareAdj(capscale(Y_repro~Dim1+Dim2+Dim3+Dim4, data=df_final,
distance="euclidean"))$adj.r.squared))
45 cat(sprintf("Model 2 adj.R2: %.4f\n",
RsquareAdj(capscale(Y_repro~Dim1+Dim2+Dim3+Dim4+HDI_PctChange, data=df_final,
distance="euclidean"))$adj.r.squared))
46 cat(sprintf("Model 3 adj.R2: %.4f\n", RsquareAdj(capscale(Y_repro~Dim1+Dim2+Dim3+Dim4+Continent,
data=df_final, distance="euclidean"))$adj.r.squared))
47 cat(sprintf("Model 4 adj.R2: %.4f\n",
RsquareAdj(capscale(Y_repro~Dim1+Dim2+Dim3+Dim4+HDI_PctChange+Continent, data=df_final,
distance="euclidean"))$adj.r.squared))
48 # --- 6. PAÏSOS PER REGIÓ ---
49 cat("\n===== \n")
50 cat("Països per regió utilitzats\n")
51 cat("===== \n")
52 print(paisos_regio)
```

Script 8. Mapamundi distribució UV i rs1801133

```
1 # Mapa mundial – Freqüència al·lèlica rs1801133 (MTHFR C677T)
2 # Dades: 1000 Genomes Project Phase 3 (26 sub-poblacions)
3 # Paquets: readxl, ggplot2, scatterpie, maps
4 library(readxl); library(ggplot2); library(scatterpie); library(maps)
5 # --- 1. CÀRREGA DE DADES ---
6 raw <- as.data.frame(read_excel(file.choose(), sheet="Freqs"))
7 # --- 2. PARSEIG DE pop_code I FREQUÈNCIES ---
8 raw$pop_code <- sub(";", "*", "", trimws(raw$Population))
9 df <- raw[raw$pop_code %in% c("ALL","AFR","AMR","EAS","EUR","SAS"), ]
10 allele_col <- as.character(df[["Allele: frequency (count)"]])
11 df$freq_G <- as.numeric(sub("G:\\s*([0-9.]+).*", "\\1", allele_col))
12 df$freq_A <- as.numeric(sub("A:\\s*([0-9.]+).*", "\\1", allele_col))
13 # --- 3. COORDENADES (posició real + posició del pastís desplaçada) ---
14 coords <- data.frame(
15   pop_code = c("ACB","ASW","ESN","GWD","LWK","MSL","YRI",
16     "CLM","MXL","PEL","PUR",
17     "CDX","CHB","CHS","JPT","KHV",
18     "CEU","FIN","GBR","IBS","TSI",
19     "BEB","GIH","ITU","PJL","STU"),
20   lon = c(-59.54,-117.15, 5.63,-15.40, 34.77,-13.23, 3.90,
21     -75.57, -99.13,-77.04,-66.59,
22     100.80, 116.39,114.06,139.69,106.63,
23     5.00, 24.94, -1.50, -3.70, 11.25,
24     90.40, 72.57, 78.49, 74.36, 80.78),
25   lat = c( 13.19, 32.72, 6.34, 13.45, 0.61, 8.49, 7.38,
26     6.25, 19.43,-12.05, 18.22,
27     22.01, 39.91, 22.54, 35.69, 10.82,
28     52.00, 60.17, 52.50, 40.42, 43.77,
29     23.73, 23.02, 17.38, 31.55, 8.73),
30   lon_pie = c(-50.00,-117.15, 15.00,-24.00, 34.77,-26.00, 3.00,
31     -75.57, -99.13,-77.04,-74.00,
32     92.00, 116.39,124.00,139.69,113.00,
33     8.00, 24.94,-18.00, -3.70, 11.25,
34     102.00, 55.00, 68.00, 62.00, 88.00),
35   lat_pie = c( 6.00, 32.72, 3.00, 19.00, 0.61, 6.00, -6.00,
36     6.25, 19.43,-12.05, 28.00,
37     18.00, 39.91, 18.00, 35.69, 3.00,
38     63.00, 60.17, 58.00, 40.42, 43.77,
39     33.00, 22.00, 8.00, 43.00, -5.00),
40   stringsAsFactors=FALSE)
41 # --- 4. UNIO I VALIDACIO ---
42 df_map <- merge(df, coords, by="pop_code")
43 df_map <- df_map[!is.na(df_map$lon) & !is.na(df_map$freq_G) & !is.na(df_map$freq_A), ]
44 if (nrow(df_map)==0) stop("df_map és buit: revisa el parseig de pop_code o freq.")
45 message("Sub-poblacions representades: ", nrow(df_map))
46 leaders <- df_map[abs(df_map$lon_pie-df_map$lon)>0.5 | abs(df_map$lat_pie-df_map$lat)>0.5, ]
47 # --- 5. GRÀFIC ---
48 world_poly <- map_data("world")
49 allele_colors <- c(freq_G="#4E84C4", freq_A="#E07B39")
50 p <- ggplot() +
51   geom_polygon(data=world_poly, aes(x=long, y=lat, group=group),
52     fill="grey88", color="white", linewidth=0.15) +
53   geom_segment(data=leaders, aes(x=lon, y=lat, xend=lon_pie, yend=lat_pie),
54     color="grey40", linewidth=0.4) +
55   geom_point(data=leaders, aes(x=lon, y=lat), color="grey20", size=1.2, shape=19) +
56   geom_scatterpie(data=df_map, aes(x=lon_pie, y=lat_pie, r=5),
57     cols=c("freq_G","freq_A"), color="white", linewidth=0.3, alpha=0.92) +
58   geom_text(data=df_map, aes(x=lon_pie, y=lat_pie-6.5, label=pop_code),
59     size=2.2, color="grey20", fontface="bold") +
60   scale_fill_manual(values=allele_colors,
61     labels=c(freq_G="Atel G (677G, referència)", freq_A="Atel A (677T, variant de risc)"),
62     name="Atel · rs1801133 (MTHFR C677T)") +
63   expand_limits(x=c(-160,160), y=c(-25,80)) +
64   coord_fixed(ratio=1.1) +
65   labs(title="Freqüència al·lèlica del polimorfisme rs1801133 (MTHFR C677T)",
66     subtitle=paste0("1000 Genomes Project Phase 3 · ", nrow(df_map), " sub-poblacions"),
67     caption="Font: Ensembl / 1000 Genomes Project Phase 3. El punt i la línia indiquen la localització real de  
cada sub-població.") +
```

```

68   theme_void(base_size=11) +
69   theme(plot.title=element_text(face="bold", size=13, margin=margin(t=10,b=4), hjust=0.5),
70         plot.subtitle=element_text(size=9.5, color="grey40", margin=margin(b=8), hjust=0.5),
71         plot.caption=element_text(size=7.5, color="grey50", margin=margin(t=6,b=6), hjust=0.5),
72         plot.margin=margin(0.5,0.5,0.5,0.5,"cm"),
73         legend.position="bottom", legend.title=element_text(size=9, face="bold"),
74         legend.text=element_text(size=8.5), legend.margin=margin(t=6))
75   # --- 6. GUARDAR ---
76   ggsave(file.path(getwd(), "Mapa_FreqAllelica_rs1801133.pdf"), p, width=20, height=10)
77   message("Desat: Mapa_FreqAllelica_rs1801133.pdf")

```

Output 1

```

--- Taula creuada: % cobertura per indicador x període ---
Indicador P1 P2 P3 P4 P5 P6 P7 Països Files
CO2      96.2 96.2 96.2 96.2 96.2 96.2 96.2 96.2 96.2
PM2.5    93.9 93.9 93.9 93.9 93.9 93.9 93.9 93.9 93.9
Occupation 86.3 86.3 86.3 86.3 86.3 86.3 86.3 86.3 86.3
Pesticides 88.2 88.2 88.2 89.2 89.2 89.2 89.6 90.1 88.7
Hidro     57.5 58.5 59.0 58.5 58.5 58.5 58.0 59.0 58.2
HDI       80.2 86.8 87.7 89.6 89.6 89.6 89.6 89.6 86.8
Poisoning 86.8 86.3 86.8 86.8 86.8 86.8 86.8 86.8 86.4
Malaria   49.1 49.1 47.6 46.2 43.4 42.0 39.6 49.1 44.4
BMI       92.9 92.9 92.9 92.9 92.9 92.9 92.9 92.9 92.9
Stillbirth 93.4 93.4 93.4 93.4 93.4 93.4 93.4 93.4 93.4
Preterm   0.0 0.0 0.0 48.1 48.1 48.1 48.1 48.1 25.2
LBW       74.1 74.1 74.1 74.1 74.1 74.1 74.1 74.1 74.1
Temp      87.3 87.3 87.3 87.3 87.3 87.3 87.3 87.3 87.3
UV        0.0 89.6 0.0 0.0 0.0 0.0 0.0 89.6 4.3
HDP       0.0 0.0 0.0 0.0 0.0 0.0 95.3 95.3 4.5
HidroM    0.0 0.0 0.0 0.0 0.0 0.0 85.8 85.8 4.1
Anomaly   87.7 87.7 87.7 87.7 87.7 87.7 87.7 87.7 87.7
Affecteds 64.6 68.4 67.5 68.4 64.6 67.9 68.4 88.7 48.7

```

Output 2

CO2	països: 204/212 files: 4284/4452	96.2%	96.2%
PM2.5	països: 204/212 files: 4284/4452	96.2%	96.2%
Occupation	països: 202/212 files: 4242/4452	95.3%	95.3%
UV	països: 190/212 files: 190/4452	4.3%	89.6%
Anomaly	països: 186/212 files: 3906/4452	87.7%	87.7%
Affecteds	països: 194/212 files: 2409/4452	54.1%	91.5%
Pesticides	països: 193/212 files: 4029/4452	90.5%	91.0%
Hidro	països: 125/212 files: 2591/4452	58.2%	59.0%
HDI	països: 190/212 files: 3865/4452	86.8%	89.6%
Poisoning	països: 184/212 files: 3864/4452	86.8%	86.8%
Malaria	països: 107/212 files: 2221/4452	49.9%	50.5%
BMI	països: 197/212 files: 4137/4452	92.9%	92.9%
Stillbirth	països: 198/212 files: 4158/4452	93.4%	93.4%
Preterm	països: 102/212 files: 1122/4452	25.2%	48.1%
LBW	països: 157/212 files: 3297/4452	74.1%	74.1%
HDP	països: 212/212 files: 202/4452	4.5%	100.0%
Temp	països: 185/212 files: 3885/4452	87.3%	87.3%
HidroM	països: 182/212 files: 182/4452	4.1%	85.8%

Figura II. Heatmap Canvis Percentuals

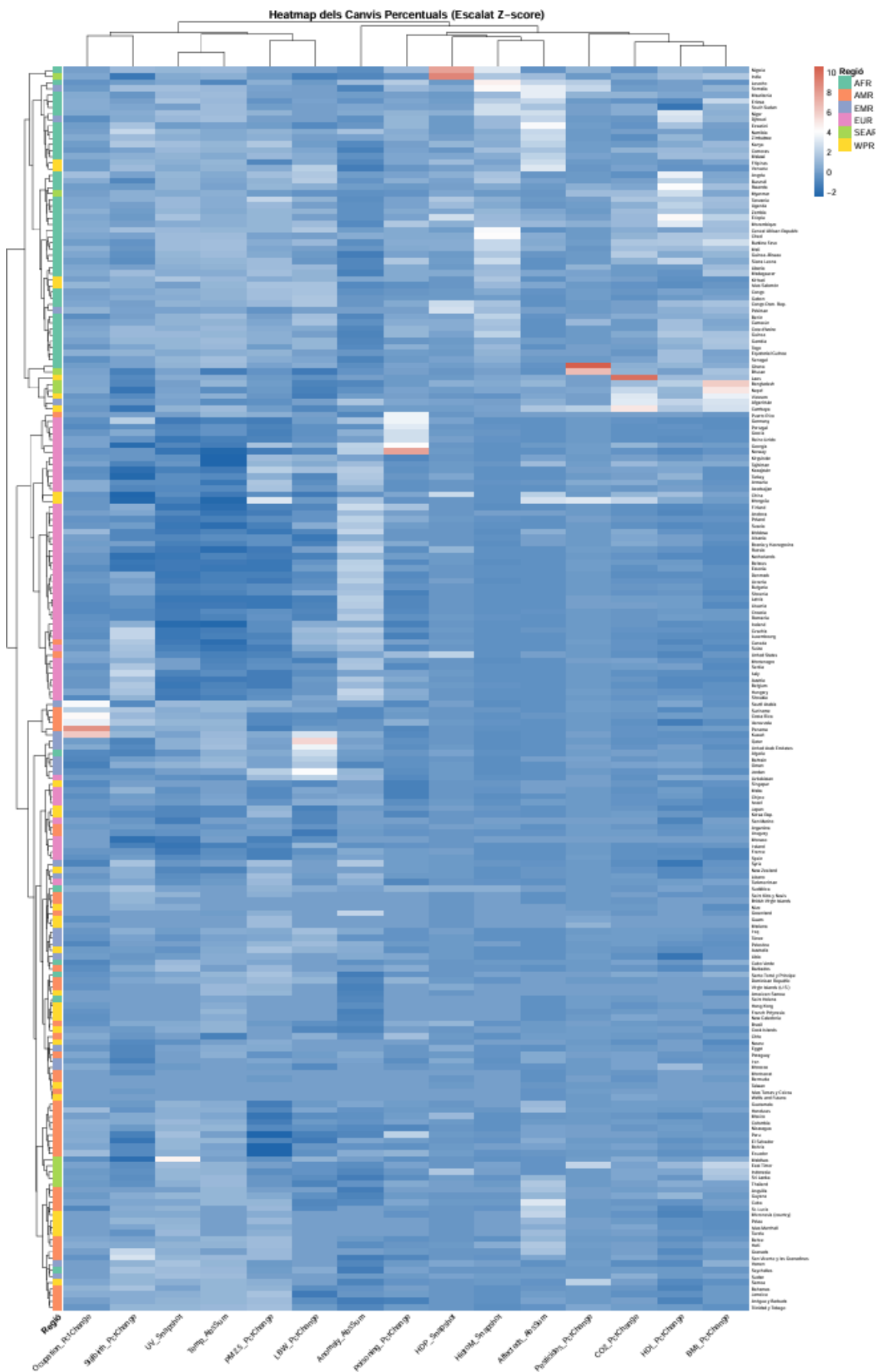
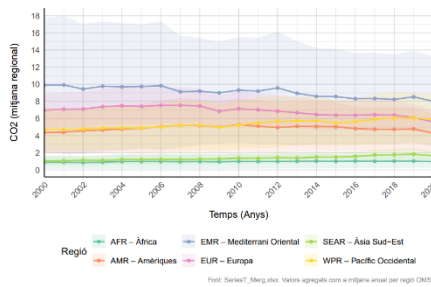
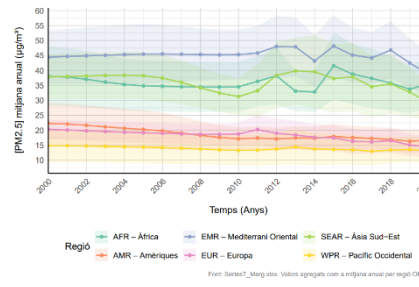


Figura 3. Sèries Temporals

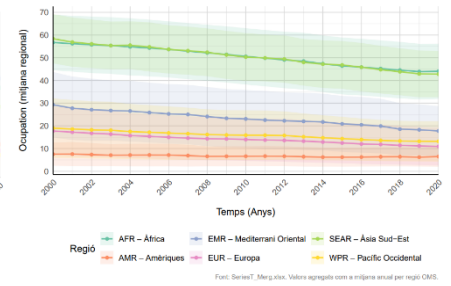
Sèrie temporal de CO₂ (2000–2020)



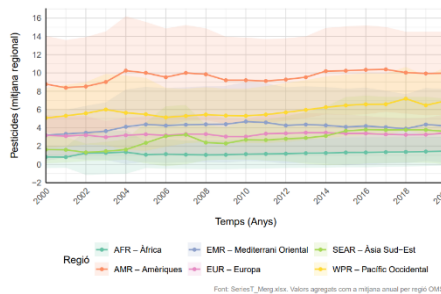
Sèrie temporal de la concentració de PM_{2.5} per continent (2000–2020)



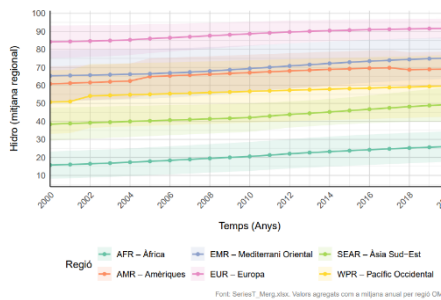
Sèrie temporal de Ocupació (2000–2020)



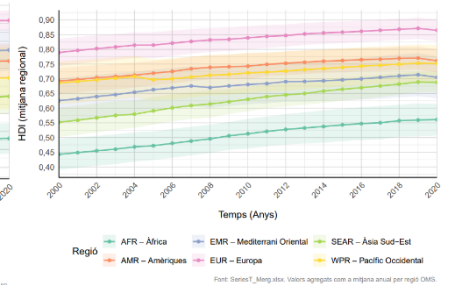
Sèrie temporal de Pesticides (2000–2020)



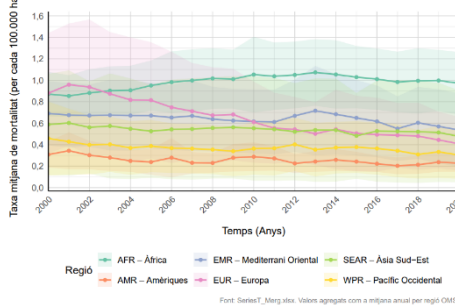
Sèrie temporal de Hidro (2000–2020)



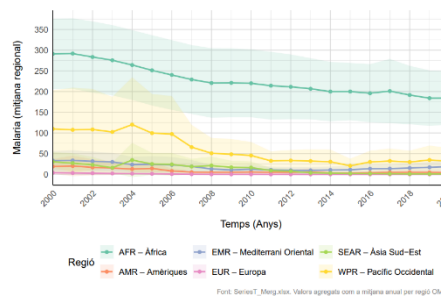
Sèrie temporal de HDI (2000–2020)



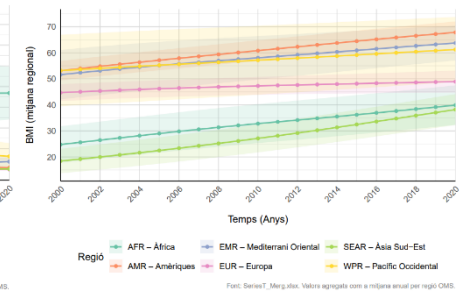
Sèrie temporal de la taxa de mortalitat per enverinament (2000–2020)



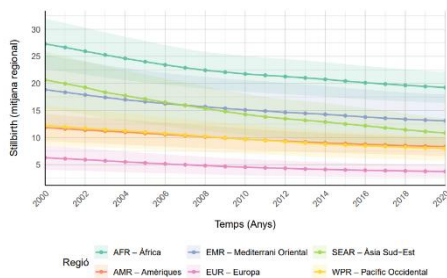
Sèrie temporal de Malaria (2000–2020)



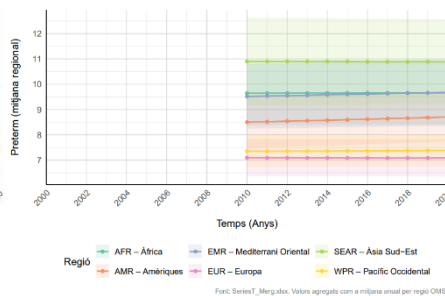
Sèrie temporal de BMI (2000–2020)



Sèrie temporal de Stillbirth (2000–2020)



Sèrie temporal de Preterm (2000–2020)



Sèrie temporal de LBW (2000–2020)

