Fluid: A Programming Language for Explorable, Transparent Research Outputs

Roly Perera

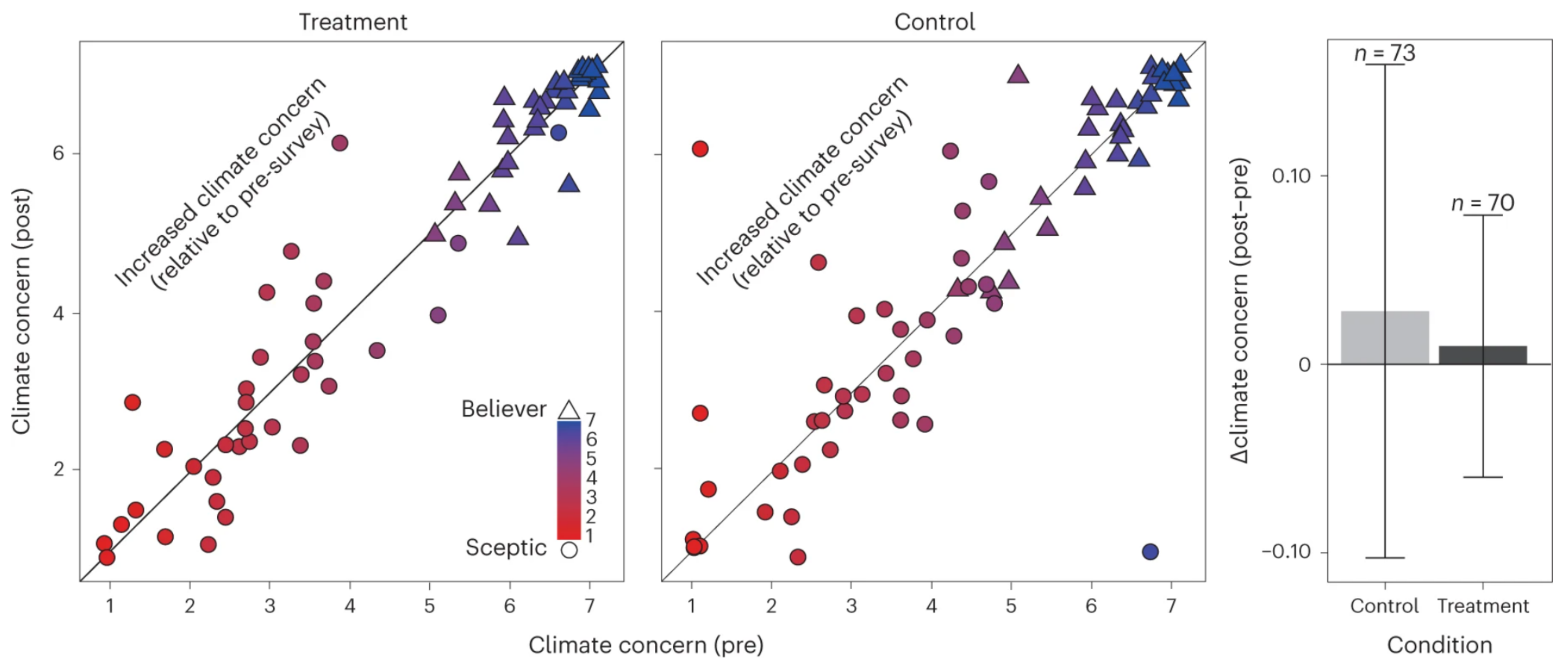
Institute of Computing for Climate Science, University of Cambridge
School of Computer Science, University of Bristol

From opaque snapshots to “open outputs”

research papers

Fig. 2: Distributions of climate beliefs before and after participating in the climate market.

From: [Participating in a climate prediction market increases concern about global warming]



Open science is increasingly the norm..

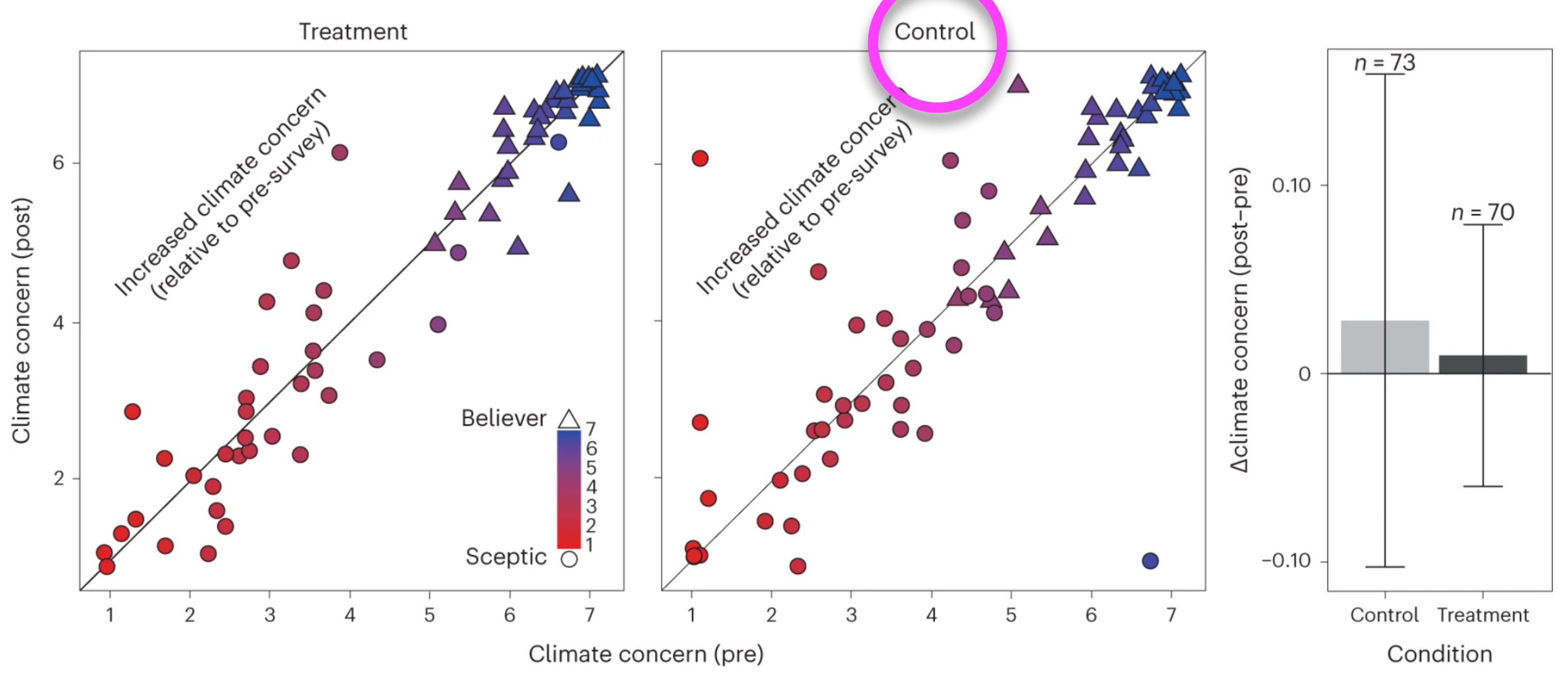
But research outputs remain **opaque**: disconnected from the data and computations used to produce them

From opaque snapshots to “open outputs”

research papers

Fig. 2: Distributions of climate beliefs before and after participating in the climate market.

From: Participating in a climate prediction market increases concern about global warming

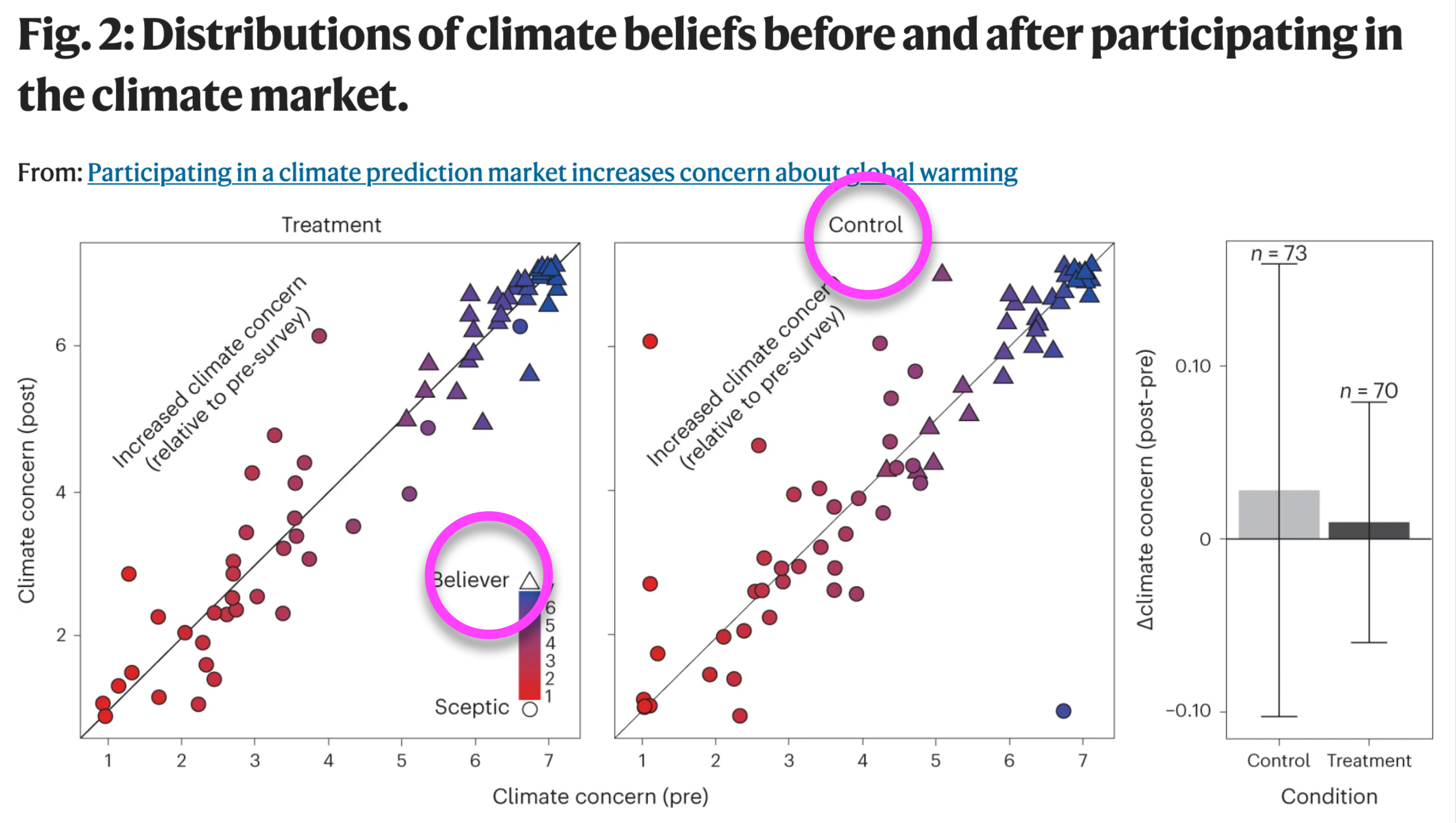


Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

From opaque snapshots to “open outputs”

research papers

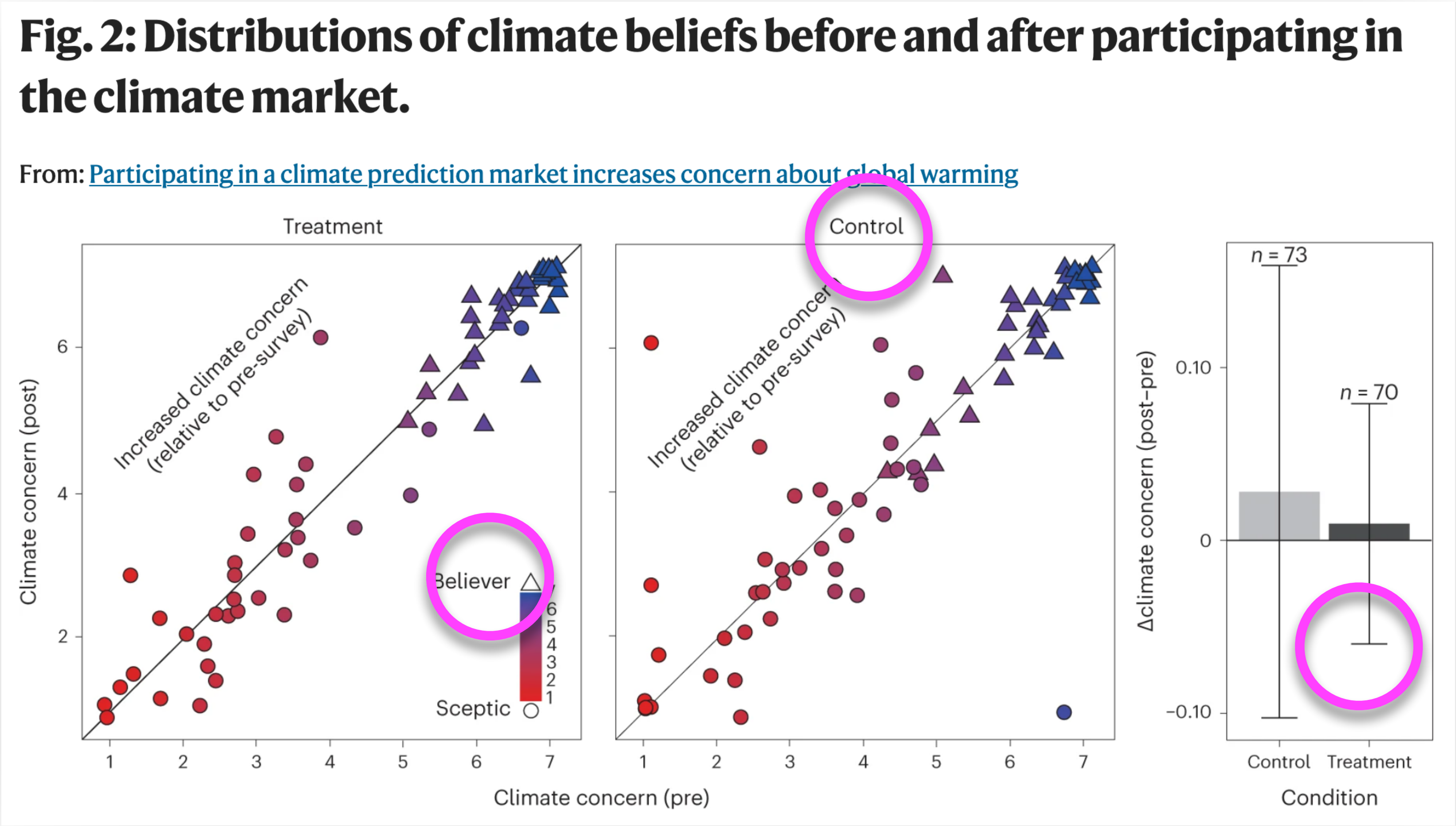


Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

From opaque snapshots to “open outputs”

research papers

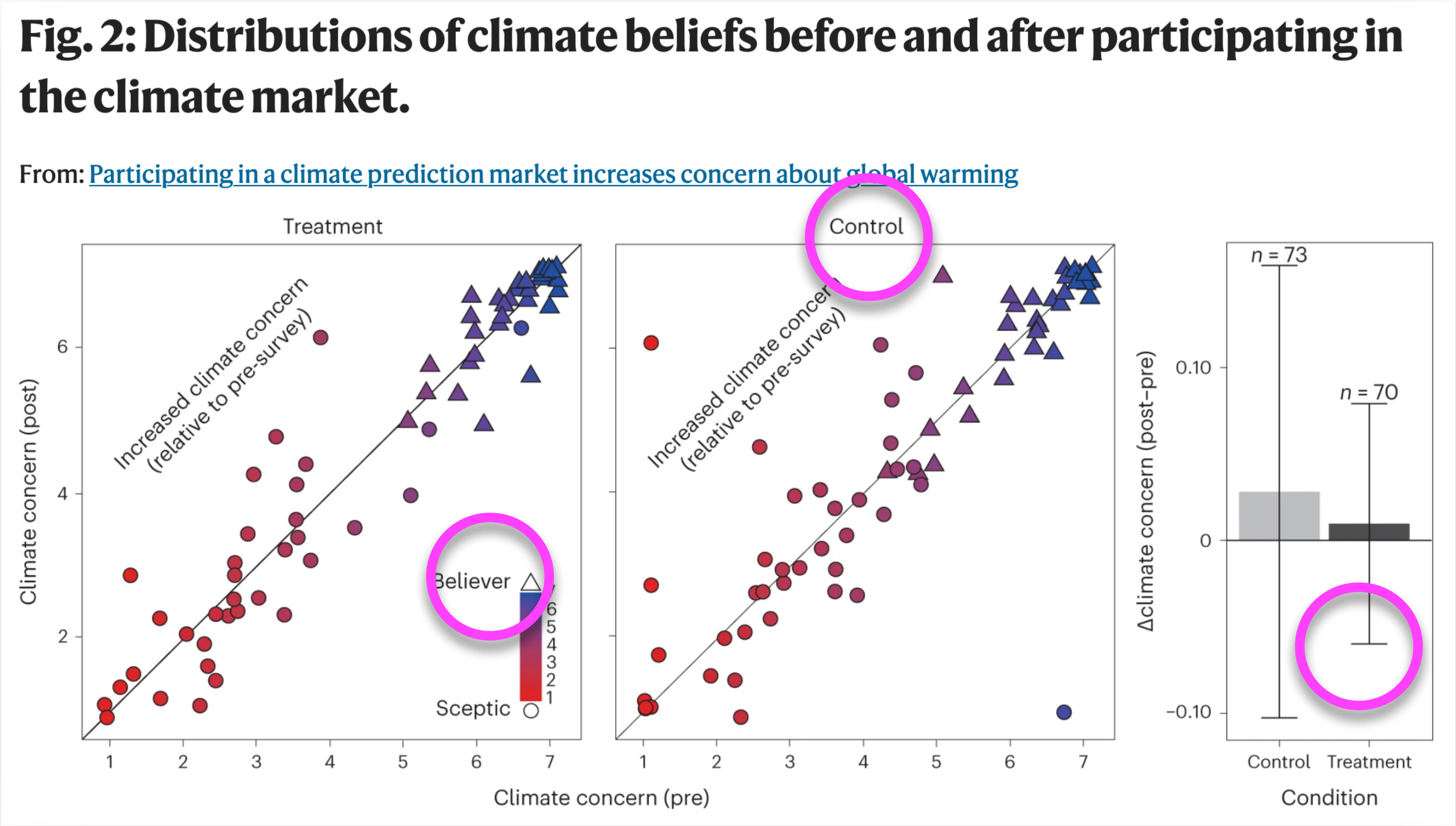


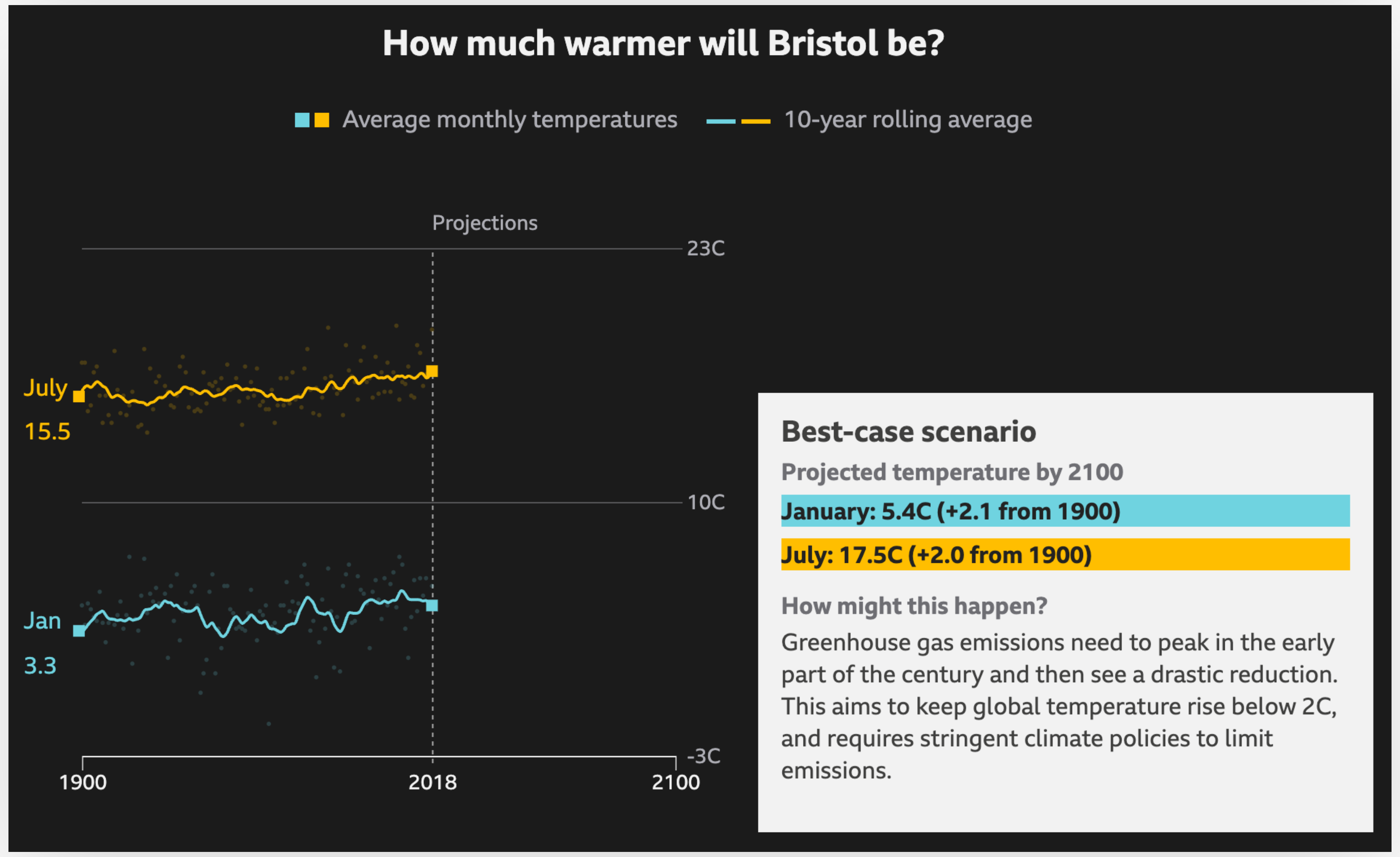
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

From opaque snapshots to “open outputs”

research papers





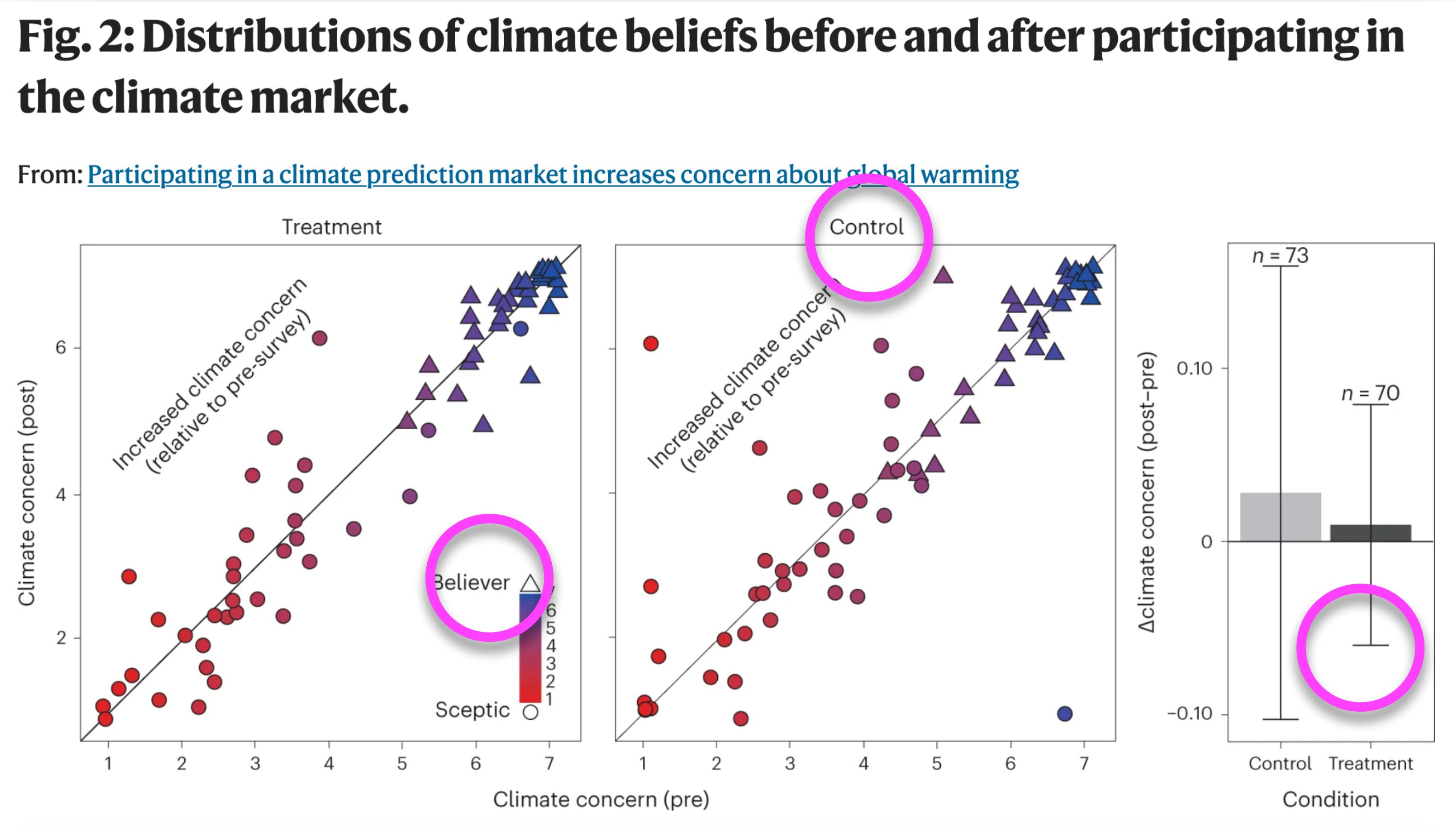
visual journalism

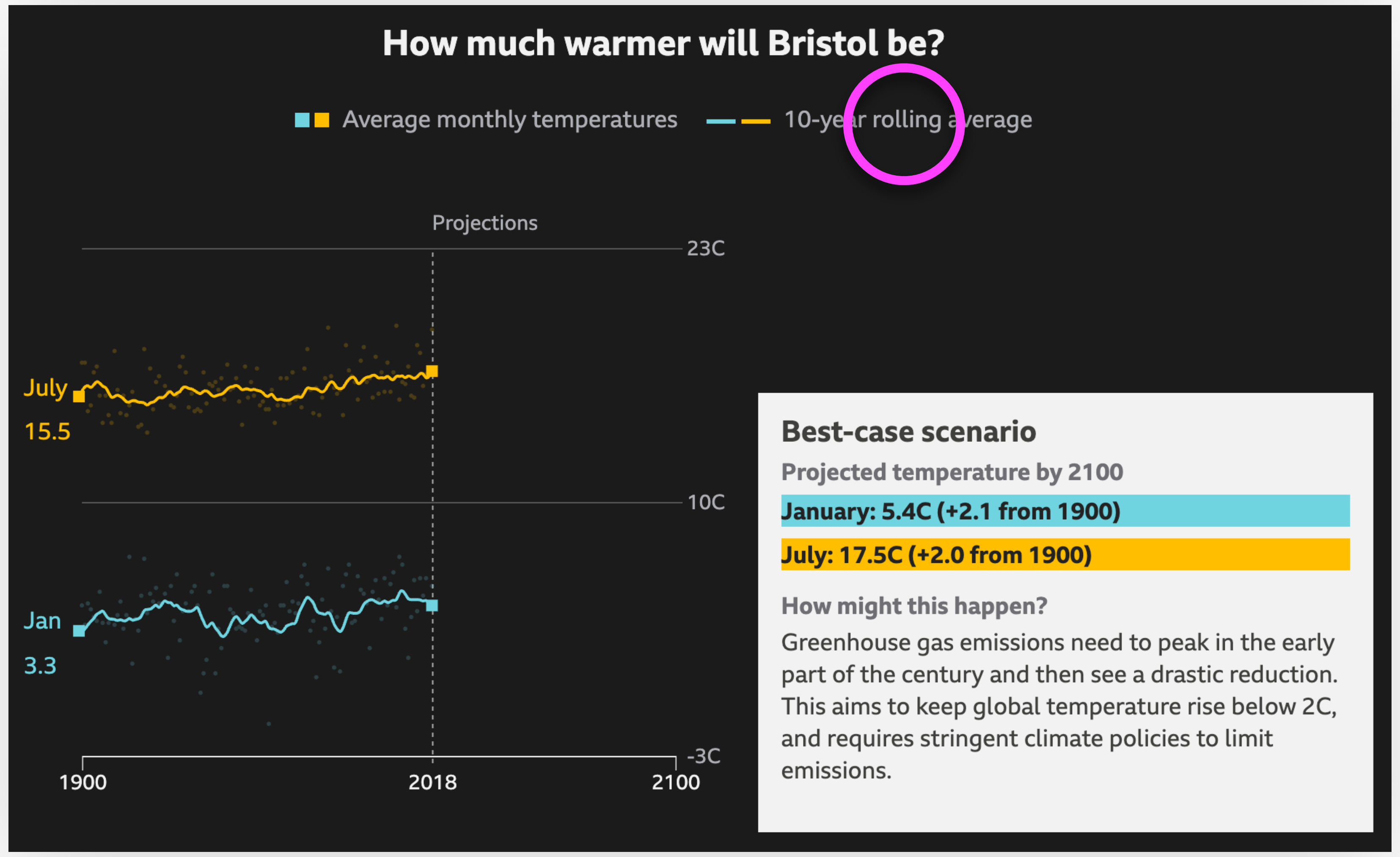
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

From opaque snapshots to “open outputs”

research papers





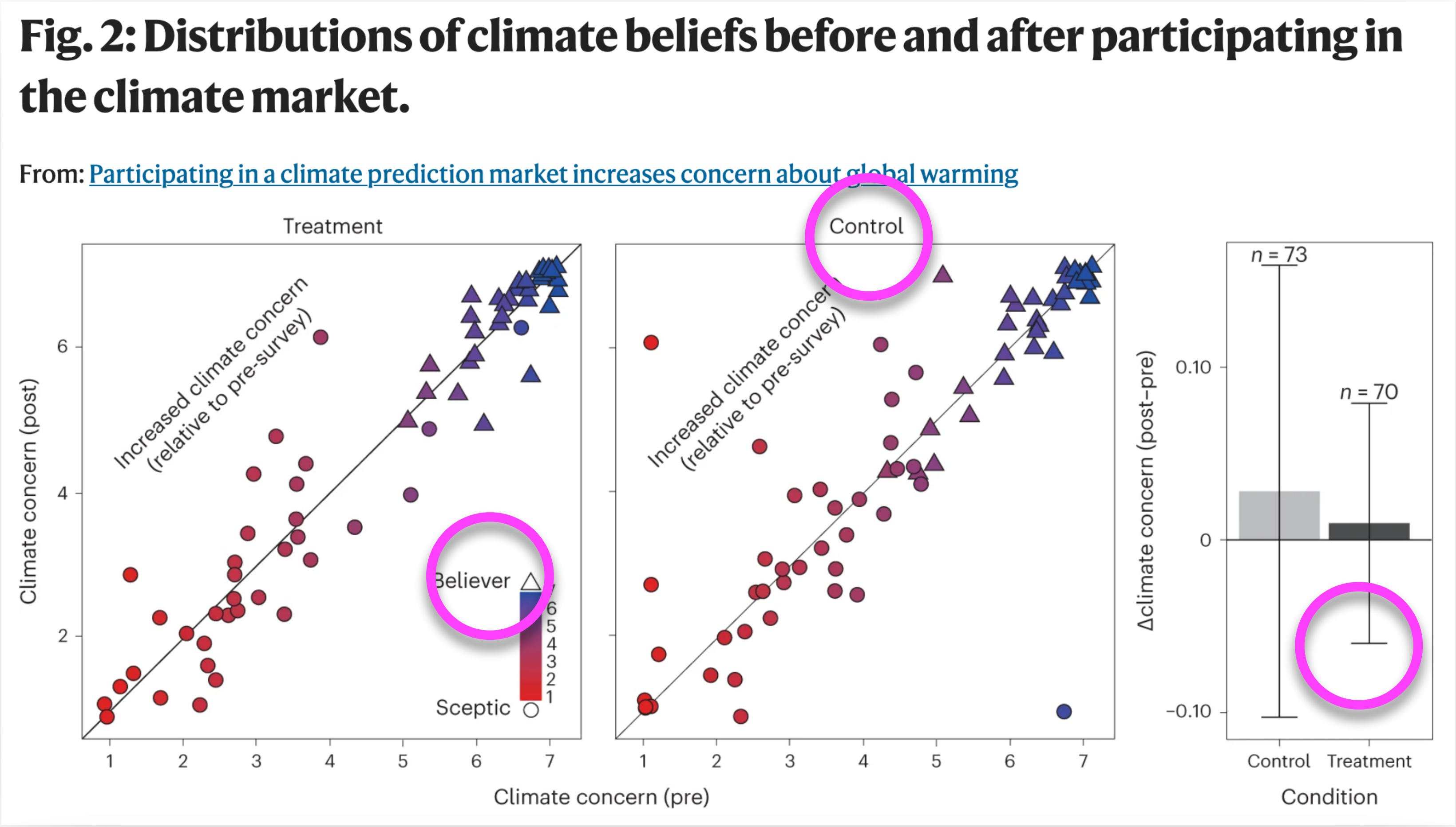
visual journalism

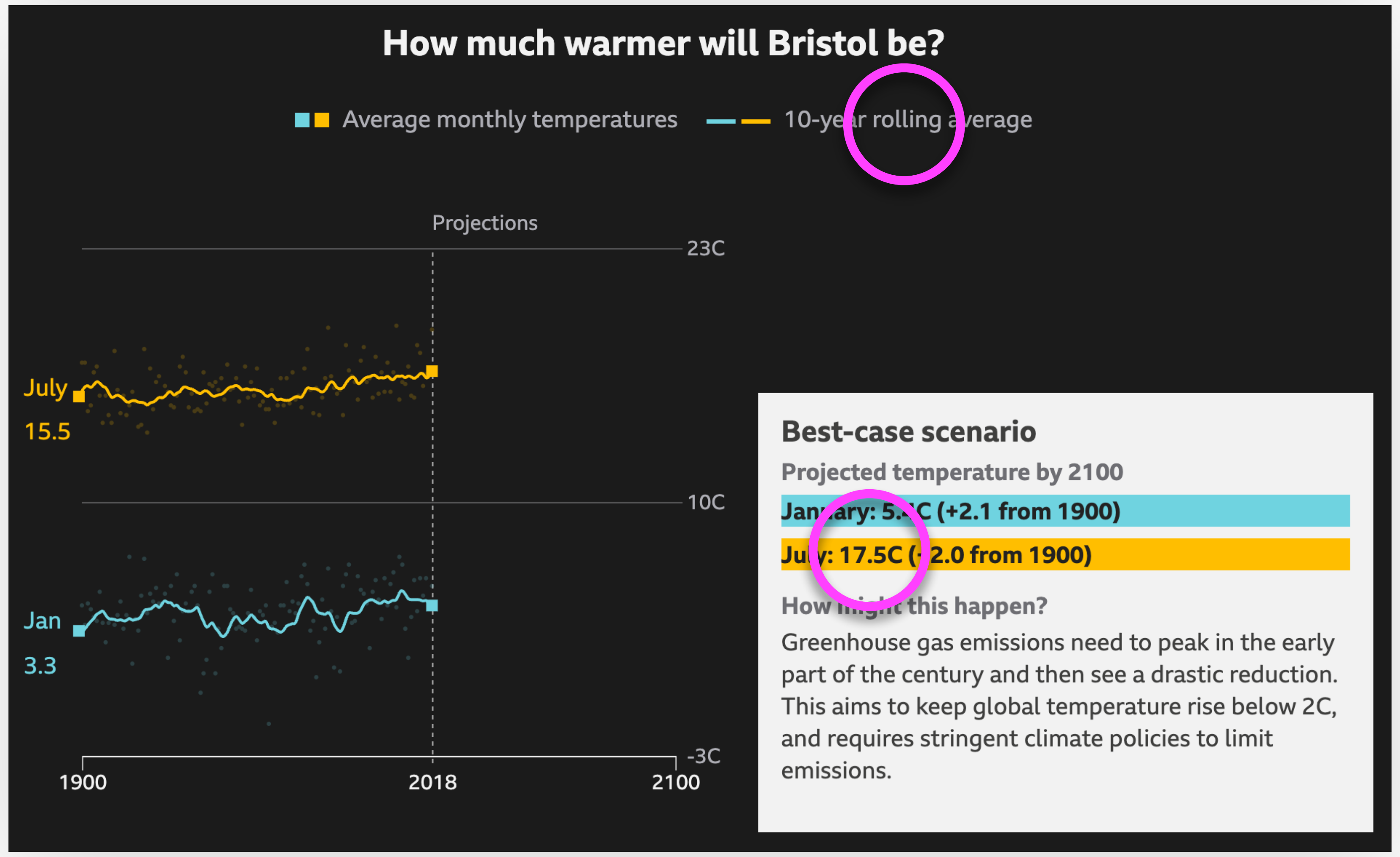
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

From opaque snapshots to “open outputs”

research papers





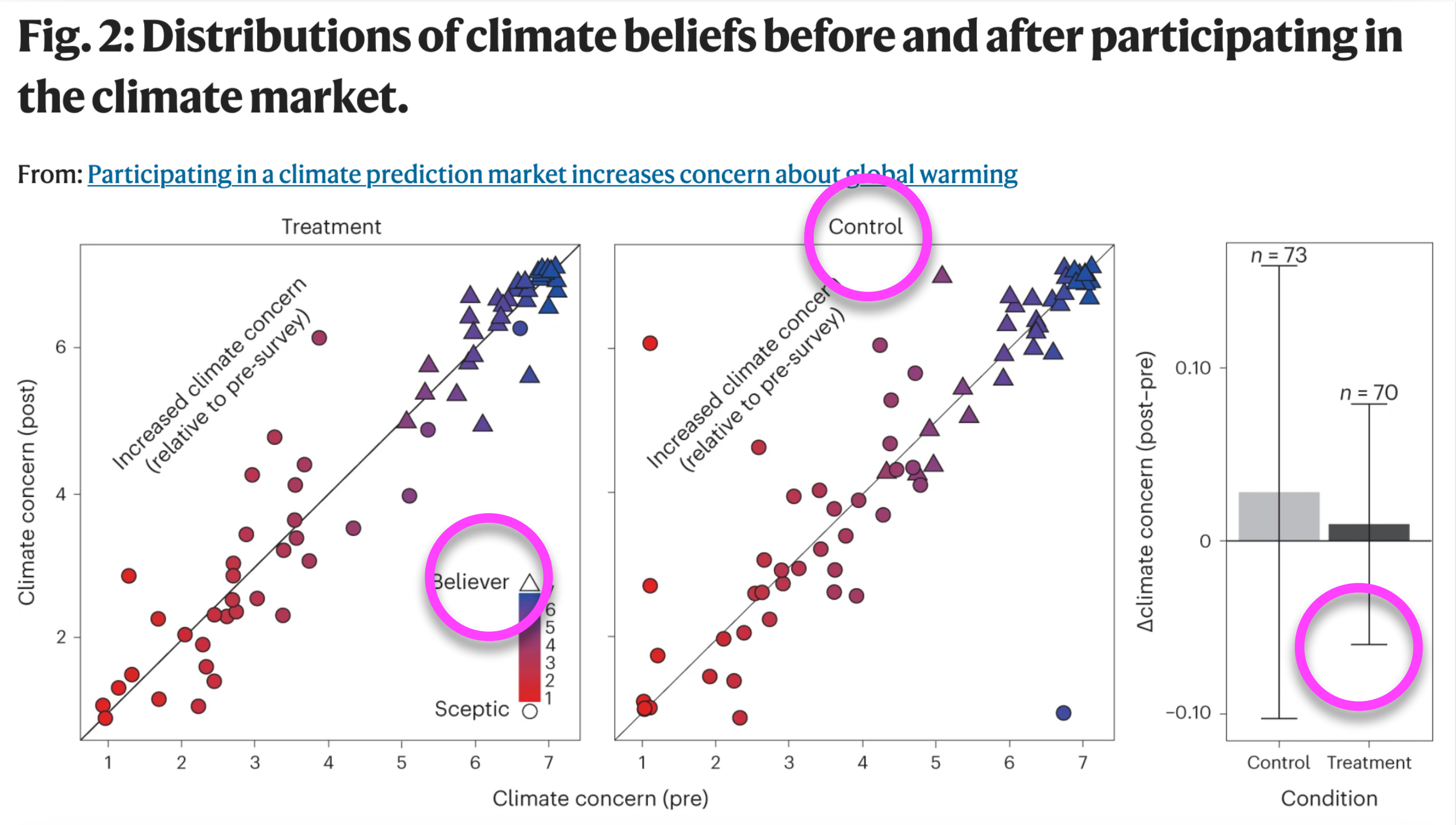
visual journalism

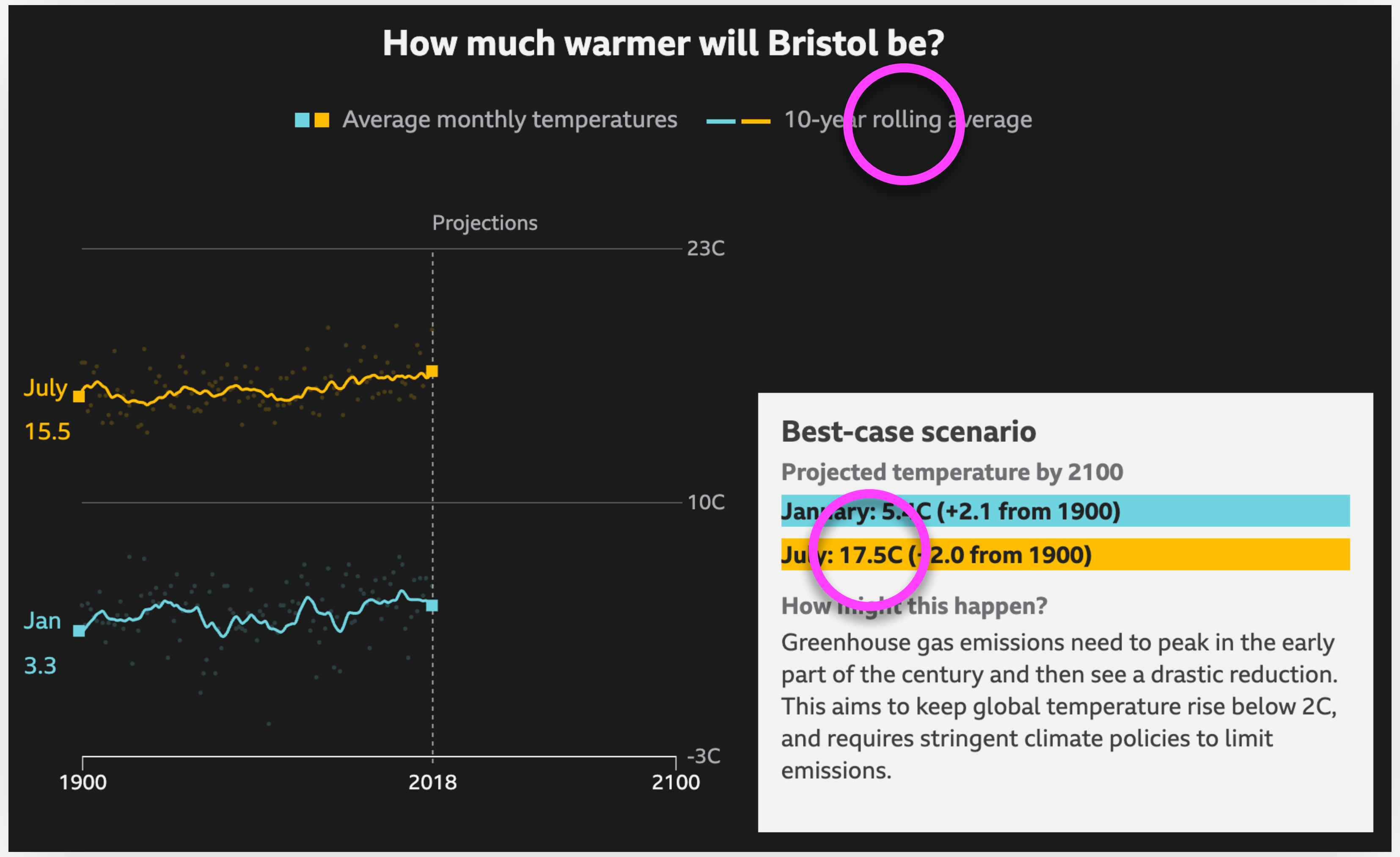
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

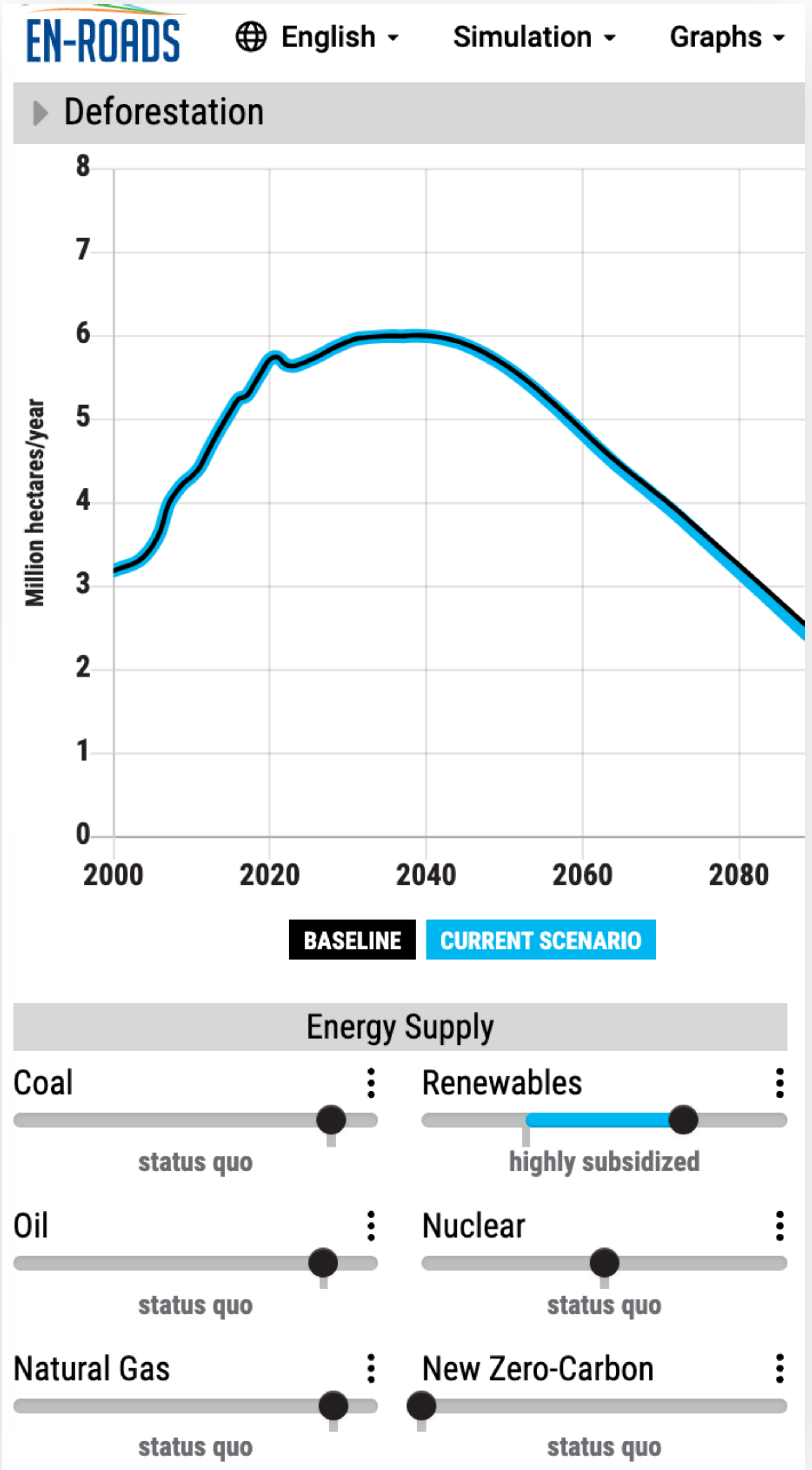
From opaque snapshots to “open outputs”

research papers





visual journalism



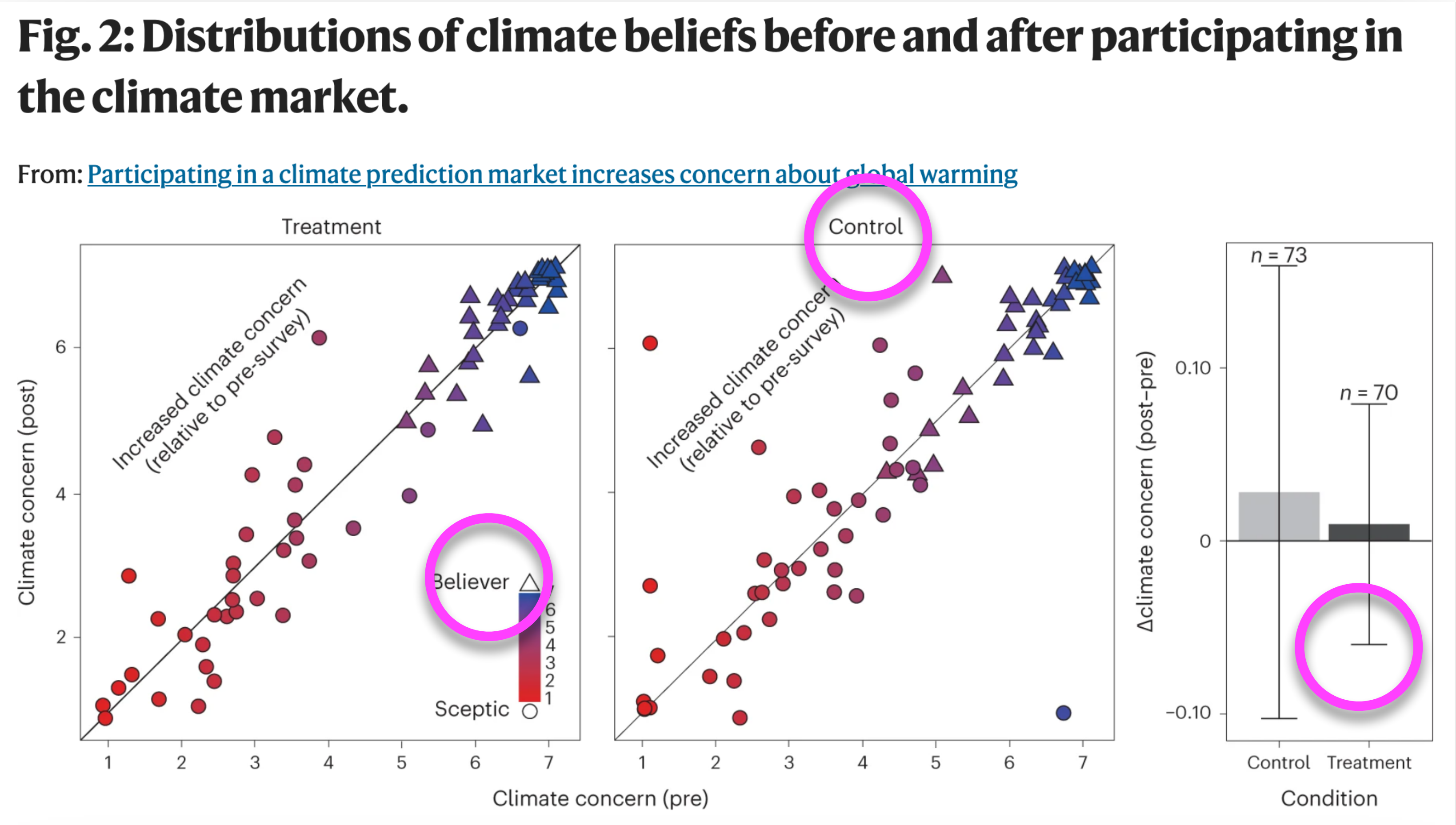
interactive simulation

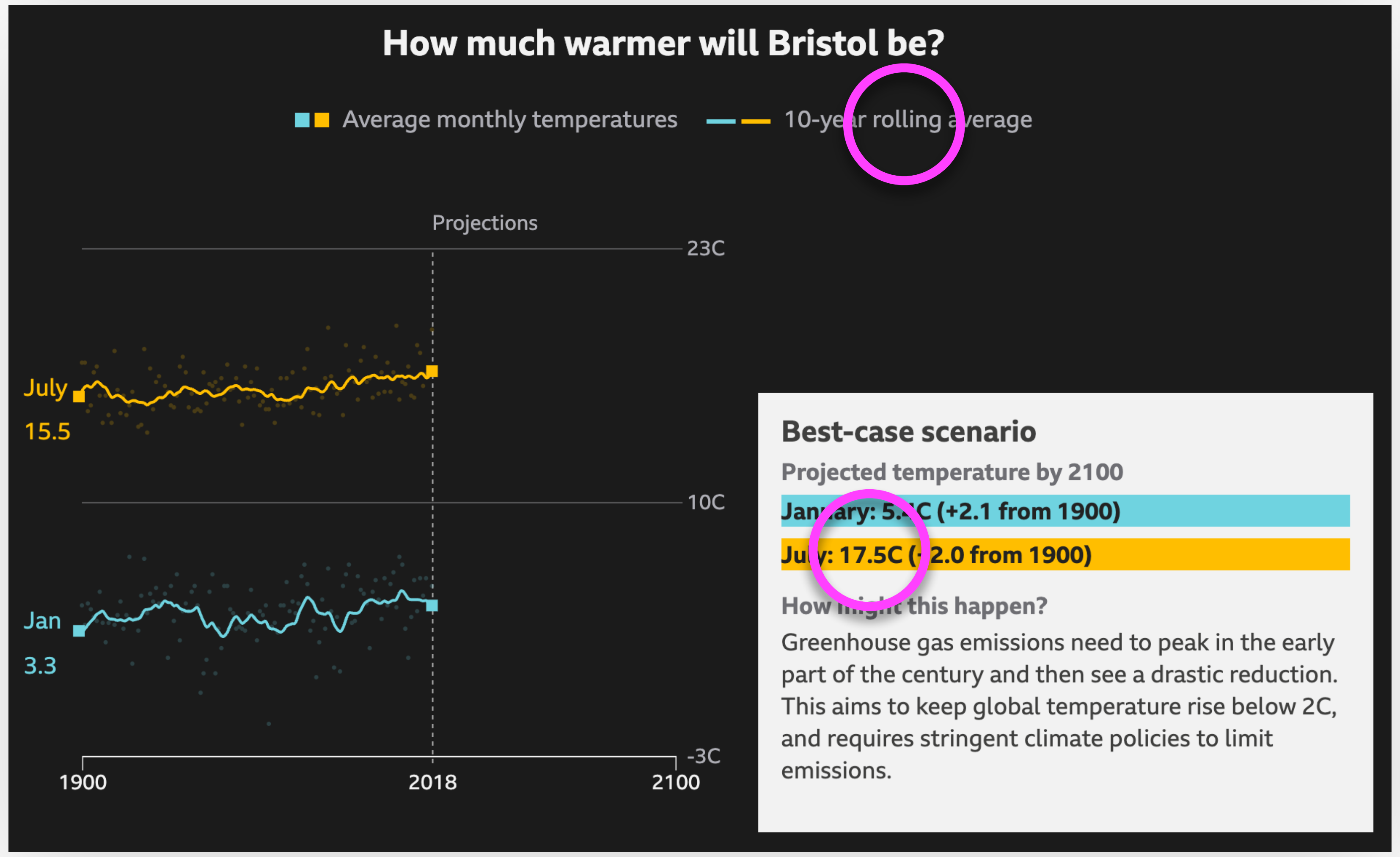
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

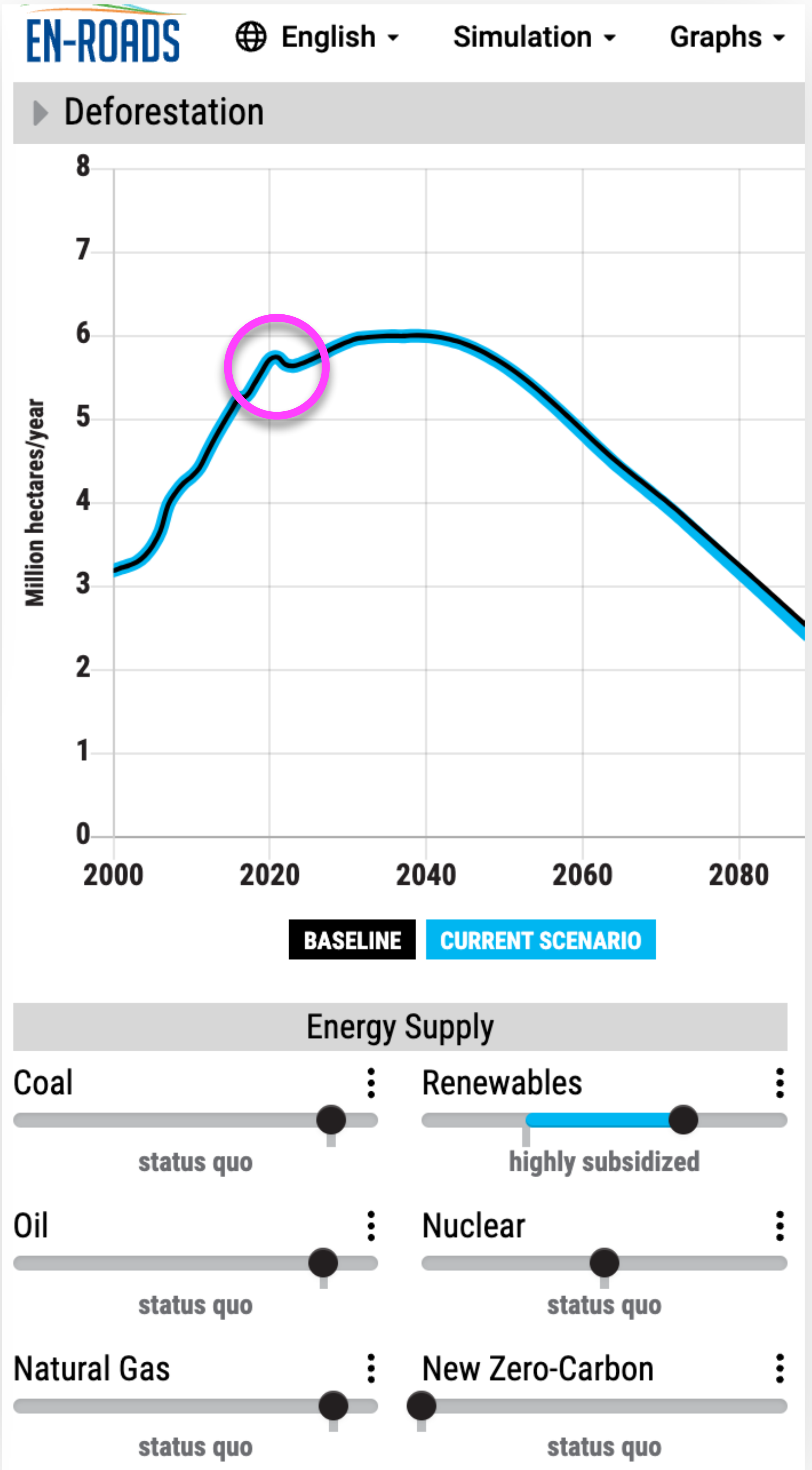
From opaque snapshots to “open outputs”

research papers





visual journalism



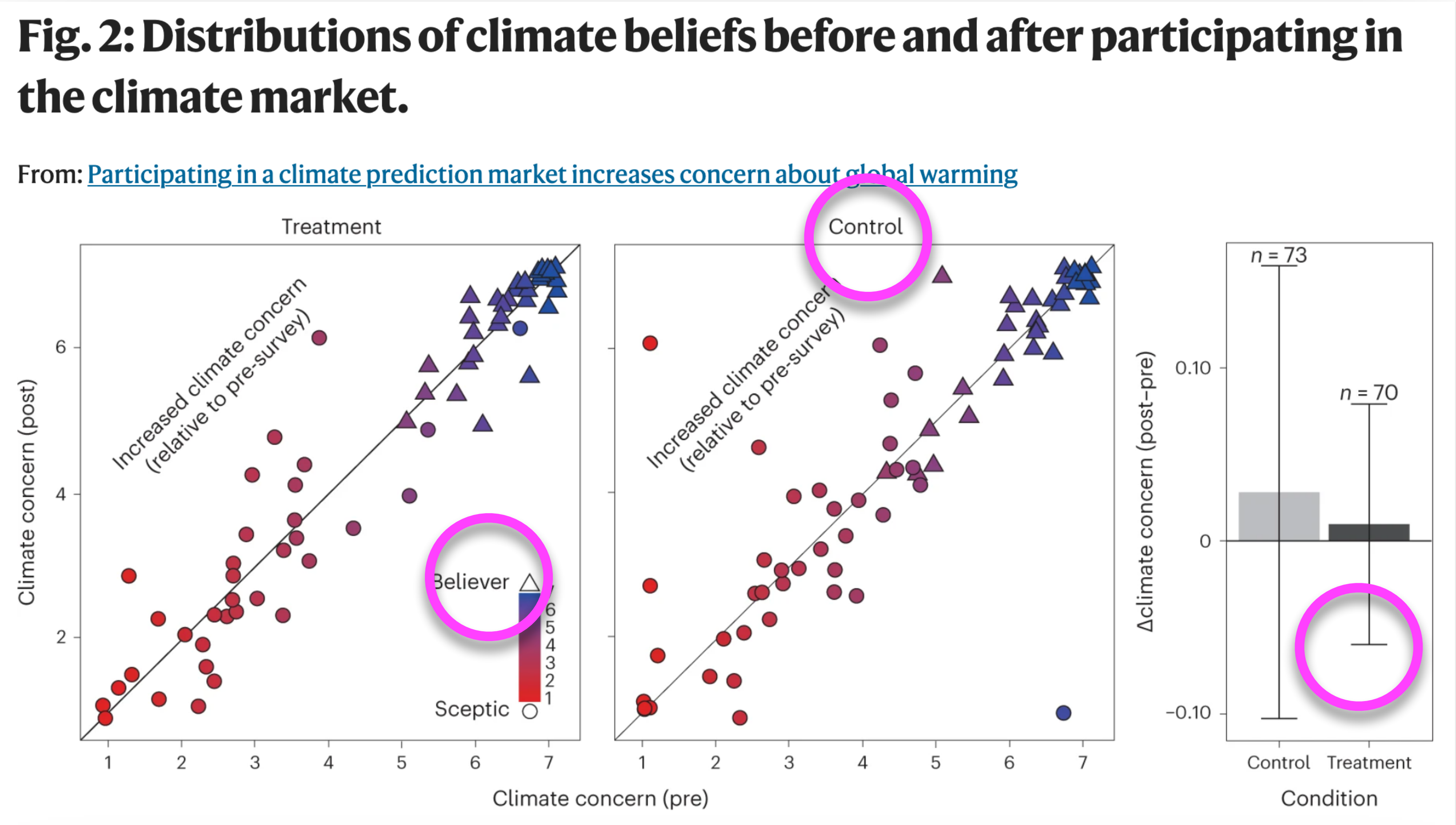
interactive simulation

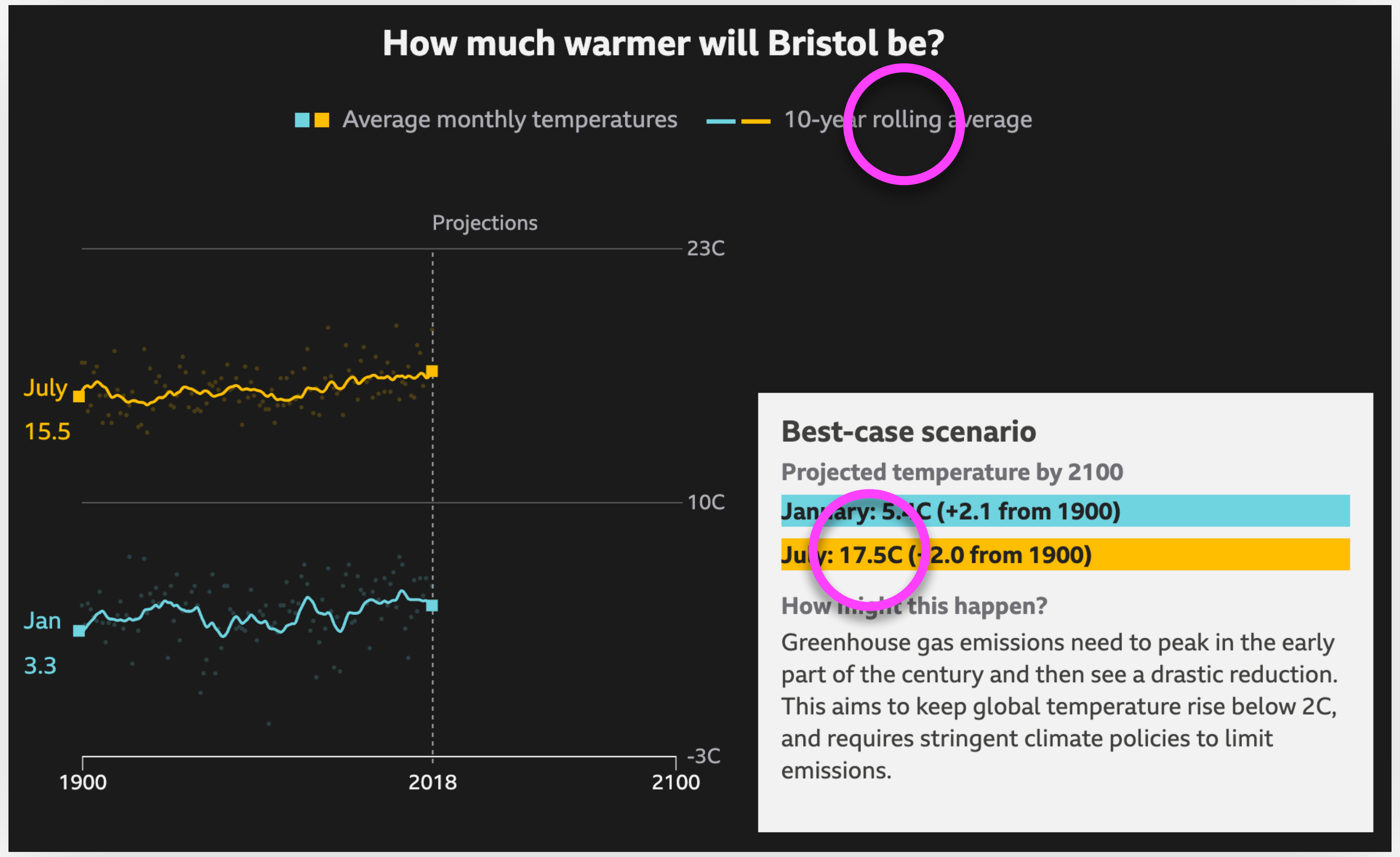
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

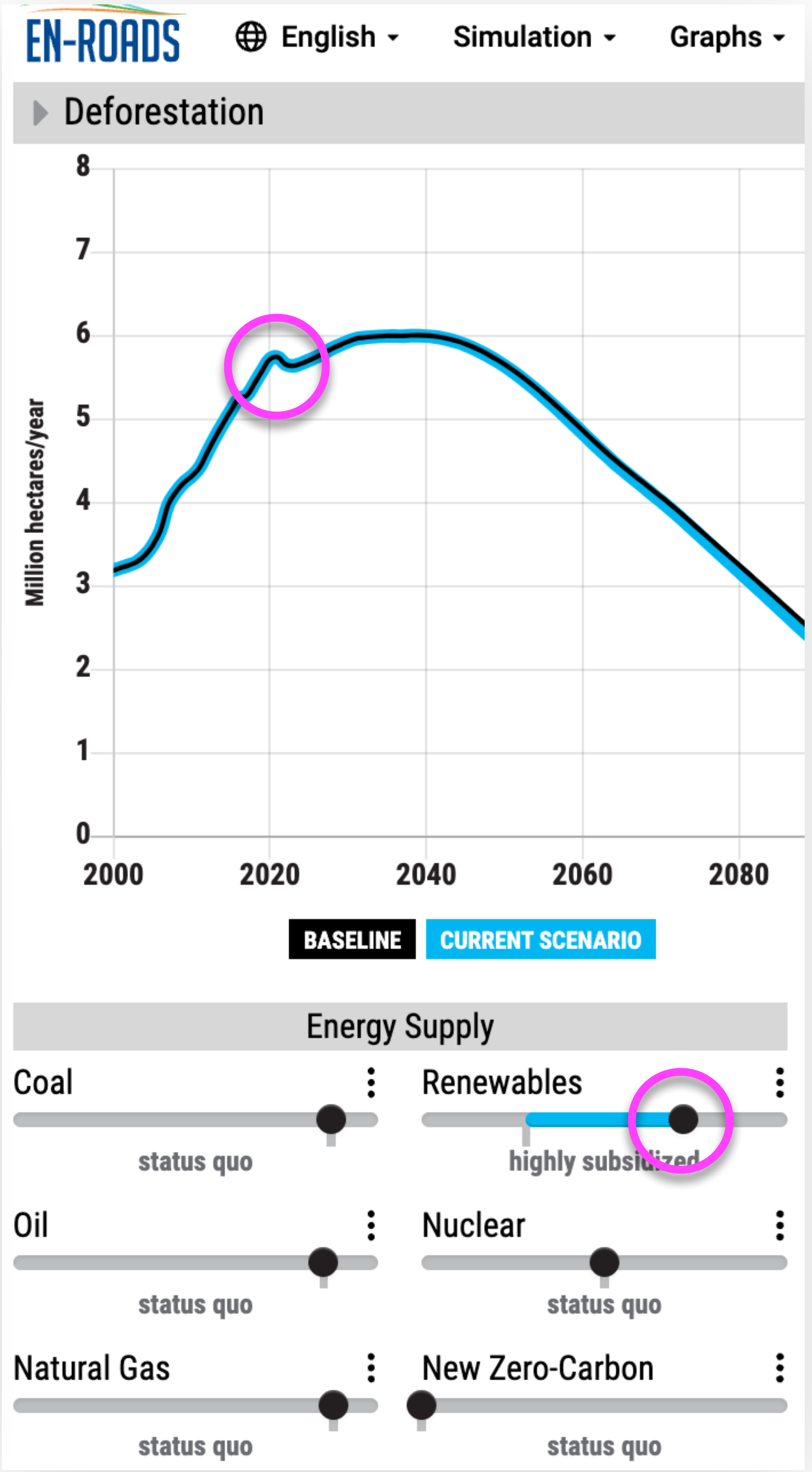
From opaque snapshots to “open outputs”

research papers





visual journalism



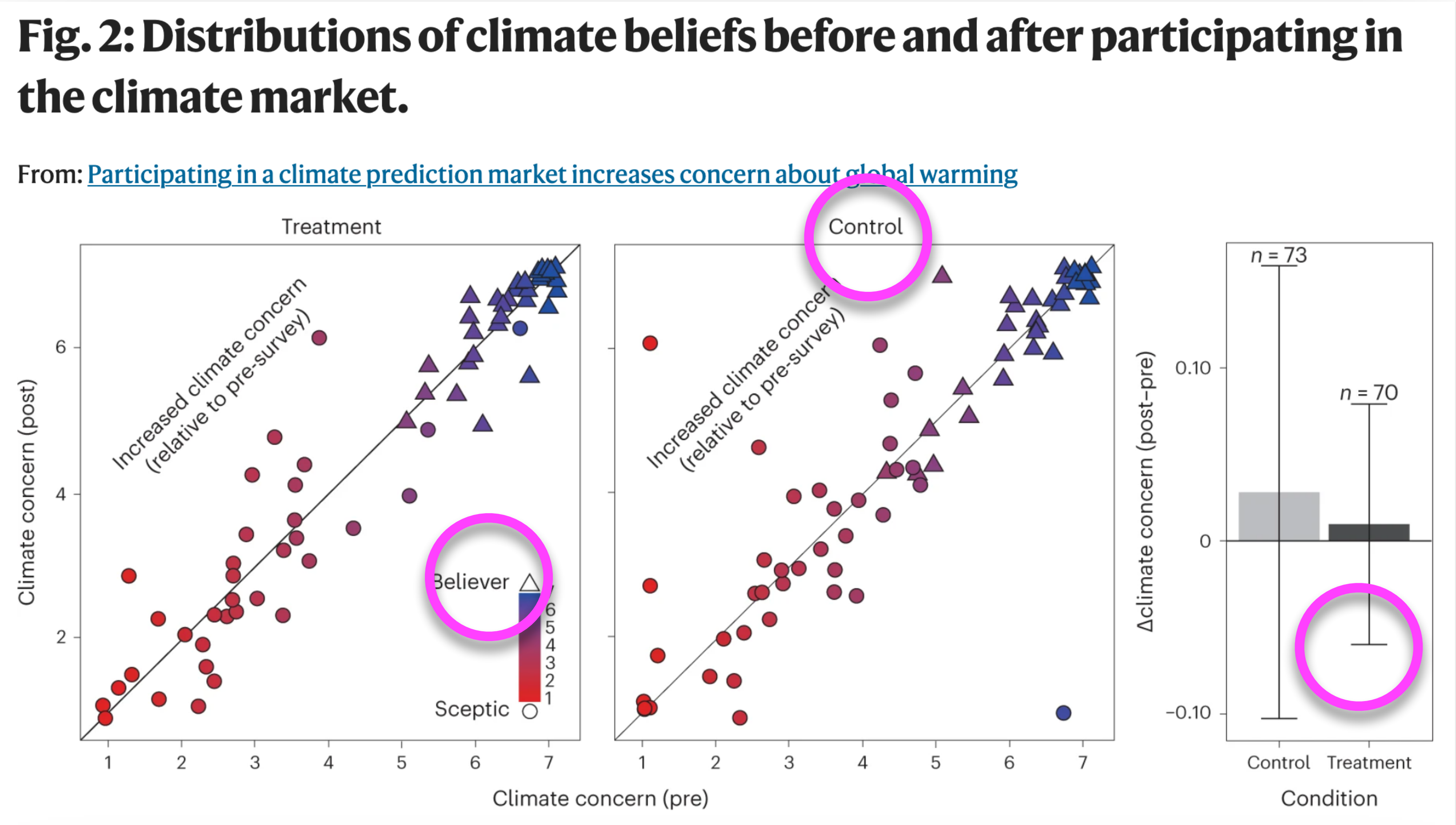
interactive simulation

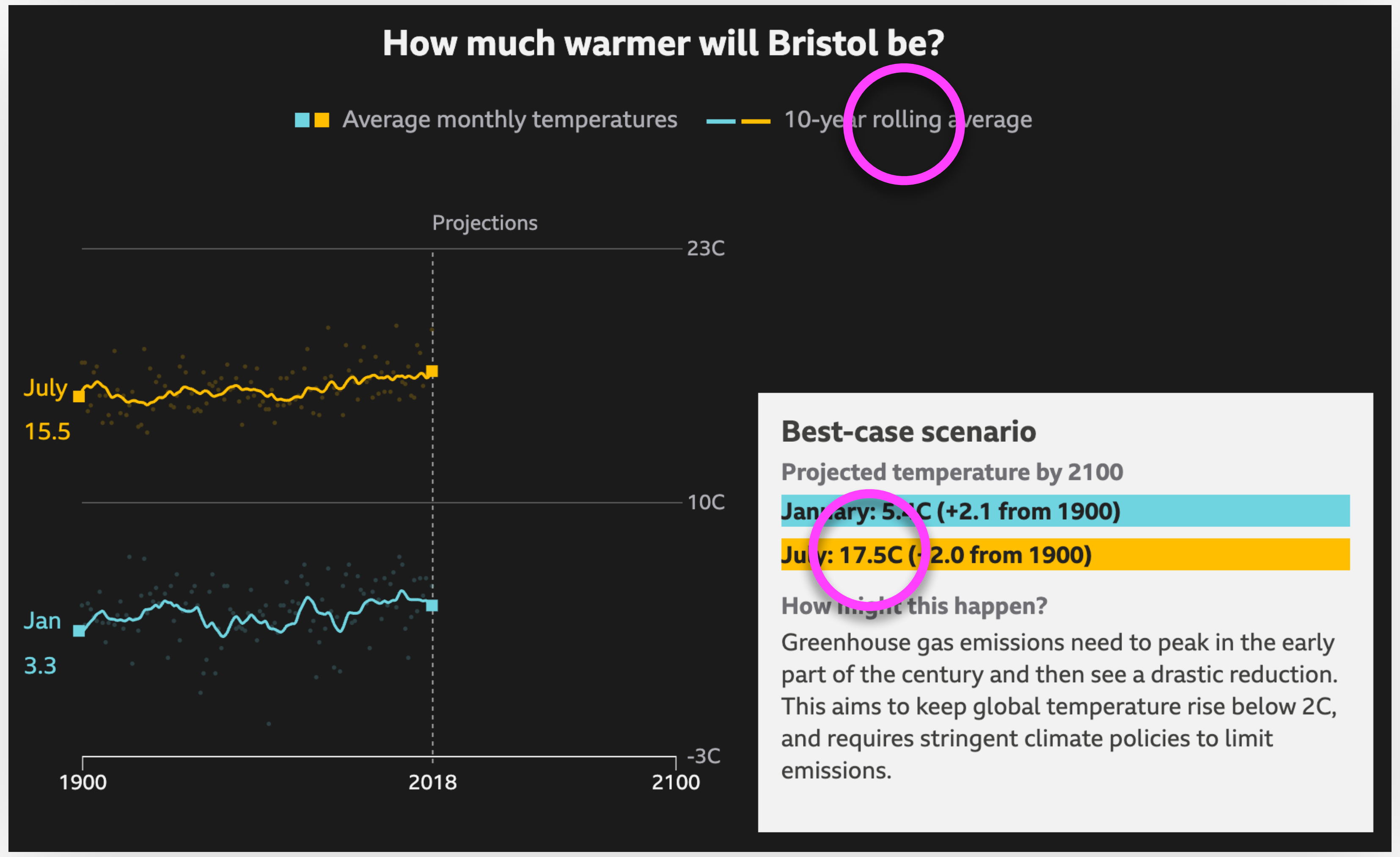
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

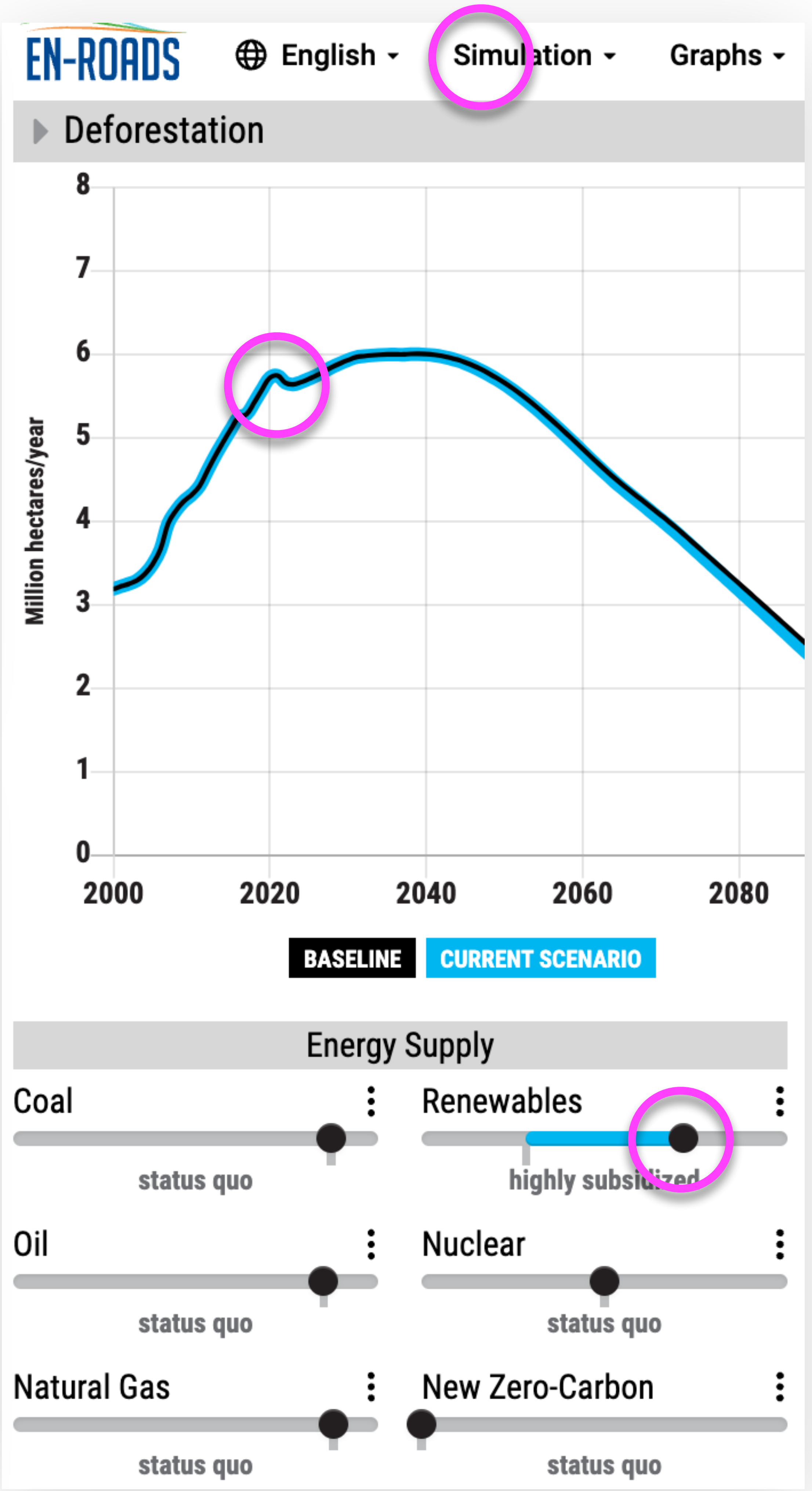
From opaque snapshots to “open outputs”

research papers





visual journalism



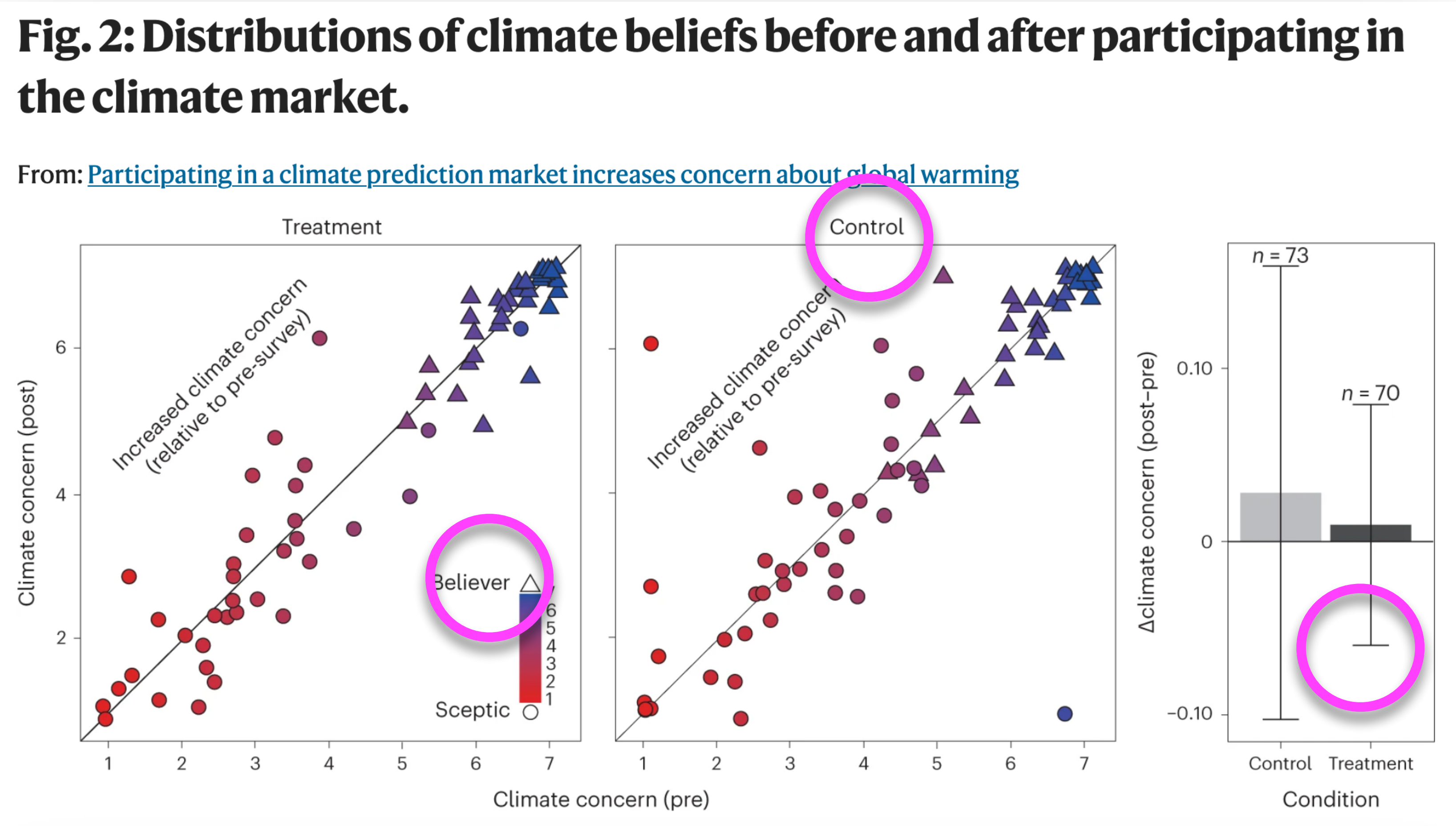
interactive simulation

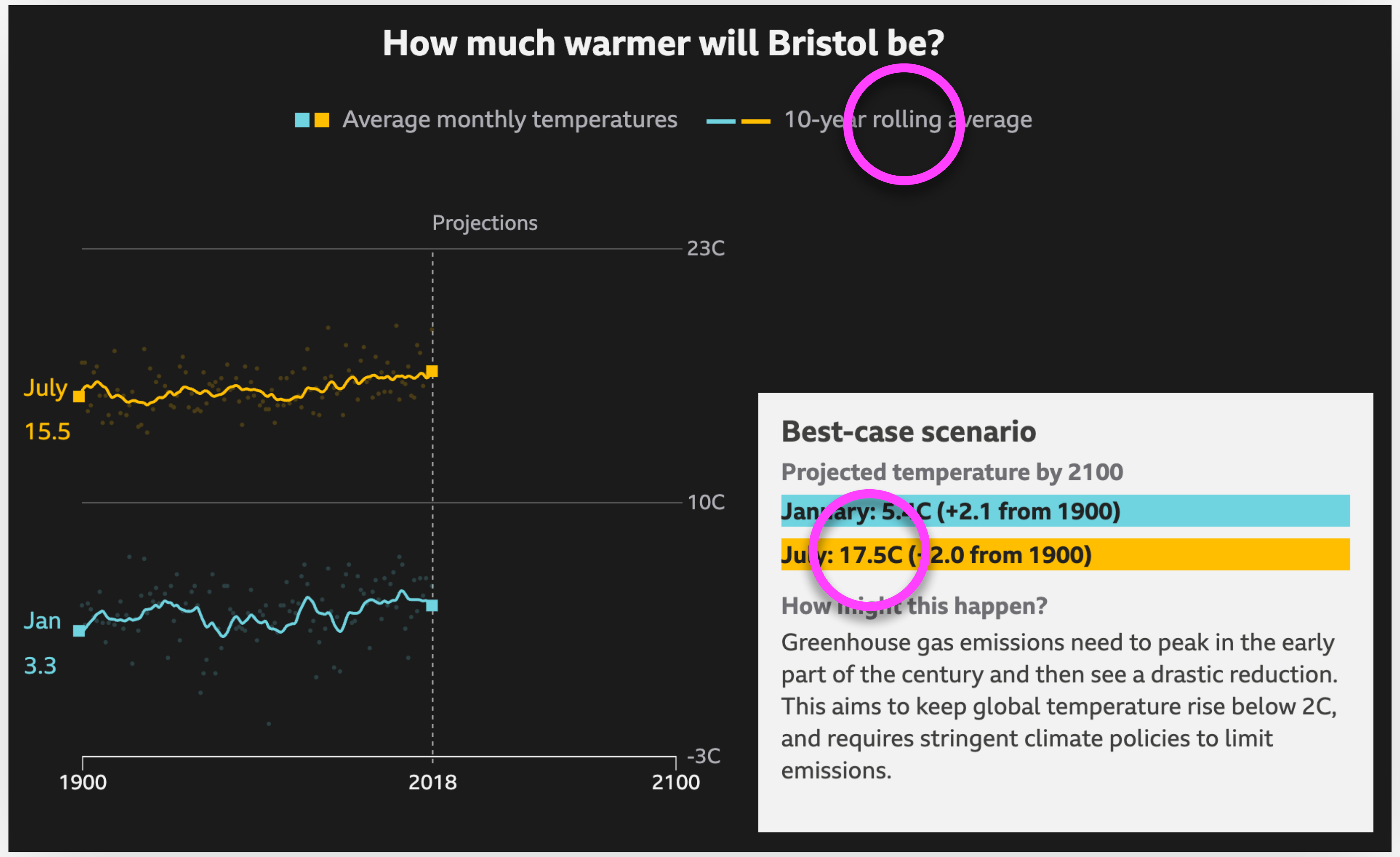
Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

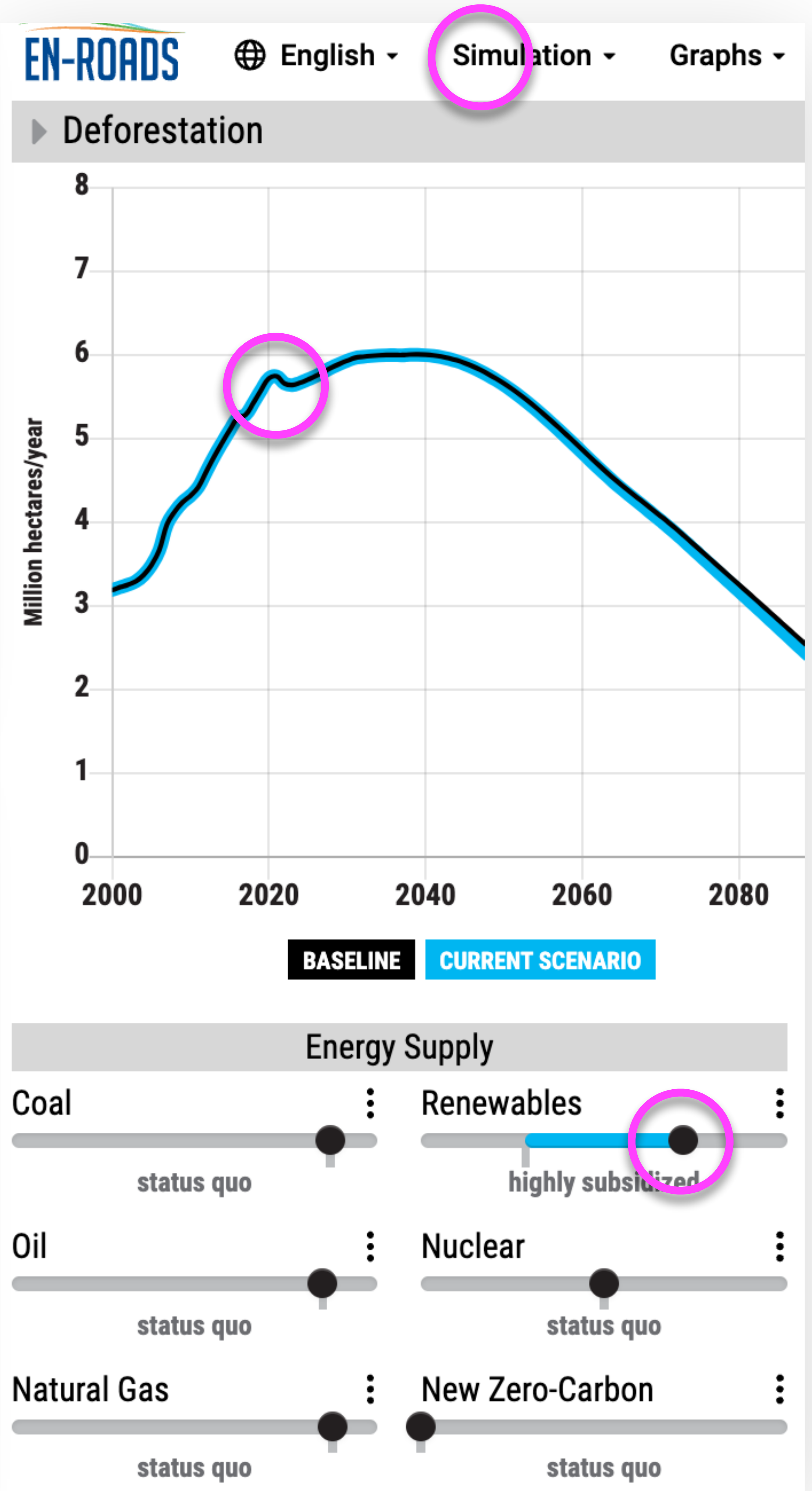
From opaque snapshots to “open outputs”

research papers





visual journalism



interactive simulation

Open science is increasingly the norm..

But research outputs remain **opaque**: disconnected from the data and computations used to produce them

What can we do to make these artifacts more **transparent and self-explanatory**?

Demo: non-renewable energy charts

```
let countries = ["BRA", "EGY", "IND", "JPN"];
let totalFor year country =
   let [ row ] = [ row | row ← nonRenewables, row.year == year, row.country == country ]
   in row.nuclearOut + row.gasOut + row.coalOut + row.petrolOut;
let stack year = [ { y: country, z: totalFor year country } | country ← countries ];
let yearData year = [ row | row ← nonRenewables, row.year == year, row.country `elem` countries ]
in MultiView {|
   "bar-chart" := BarChart {
      caption: "Non-renewables output",
      size: { width: 275, height: 185 },
      stackedBars: [ { x: numToStr year, bars: stack year } | year ← [2014..2018] ]
   },
   "scatter-plot" := ScatterPlot {
      caption: "",
      points: [ {
         x: sum [ row.nuclearOut | row ← yearData year ],
         y: sum [ row.nuclearCap | row ← yearData year ]
      } | year ← [2014..2018] ],
      xlabel: "Nuclear capacity",
      ylabel: "Nuclear output"
   }
|}
```

Demo: non-renewable energy charts

```
let countries = ["BRA", "EGY", "IND", "JPN"];
let totalFor year country =
   let [ row ] = [ row | row ← nonRenewables, row.year == year, row.country == 
   in row.nuclearOut + row.gasOut + row.coalOut + row.petrolOut;
let stack year = [ { y: country, z: totalFor year country } | country ← countrie
let yearData year = [ row | row ← nonRenewables, row.year == year, row.country `elem` countries ]
in MultiView {|
   "bar-chart" := BarChart {
      caption: "Non-renewables output",
      size: { width: 275, height: 185 },
      stackedBars: [ { x: numToStr year, bars: stack year } | year ← [2014..201
   },
   "scatter-plot" := ScatterPlot {
      caption: "",
      points: [ {
         x: sum [ row.nuclearOut | row ← yearData year ],
         y: sum [ row.nuclearCap | row ← yearData year ]
      } | year ← [2014..2018] ],
      xlabel: "Nuclear capacity",
      ylabel: "Nuclear output"
   }
|}
```

Programmer describes how to map data to visual elements

Runtime analyses dependencies and provides interactions

User formulates queries by interacting with output

Demo: convolution

```
let zero n = const n;
    wrap n n_max = ((n - 1) `mod` n_max) + 1;
    extend n = min (max n 1);

let convolve image kernel method =
    let ((m, n), (i, j)) = (dims image, dims kernel);
        (half_i, half_j) = (i `quot` 2, j `quot` 2);
        area = i * j
    in  [| let weightedSum = sum [
           image!(x, y) * kernel!(i' + 1, j' + 1)
           | (i', j') ← range (0, 0) (i - 1, j - 1),
             let x = method (m' + i' - half_i) m,
             let y = method (n' + j' - half_j) n,
             x ⩾ 1, x ⩽ m, y ⩾ 1, y ⩽ n
         ] in weightedSum `quot` area
         | (m', n') in (m, n) |];
```

Demo: convolution

```
let zero n = const n;
    wrap n n_max = ((n - 1) `mod` n_max) + 1;
    extend n = min (max n 1);

let convolve image kernel method =
    let ((m, n), (i, j)) = (dims image, dims kernel);
        (half_i, half_j) = (i `quot` 2, j `quot` 2);
        area = i * j
    in  [| let weightedSum = sum [
           image!(x, y) * kernel!(i' + 1, j' + 1)
           | (i', j') ← range (0, 0) (i - 1, j - 1),
             let x = method (m' + i' - half_i) m,
             let y = method (n' + j' - half_j) n,
             x ⩾ 1, x ⩽ m, y ⩾ 1, y ⩽ n
         ] in weightedSum `quot` area
         | (m', n') in (m, n) |];
```

Programmer implements convolution in a conventional way

↓

Runtime provides interactions that reveal behaviour of convolution

↓

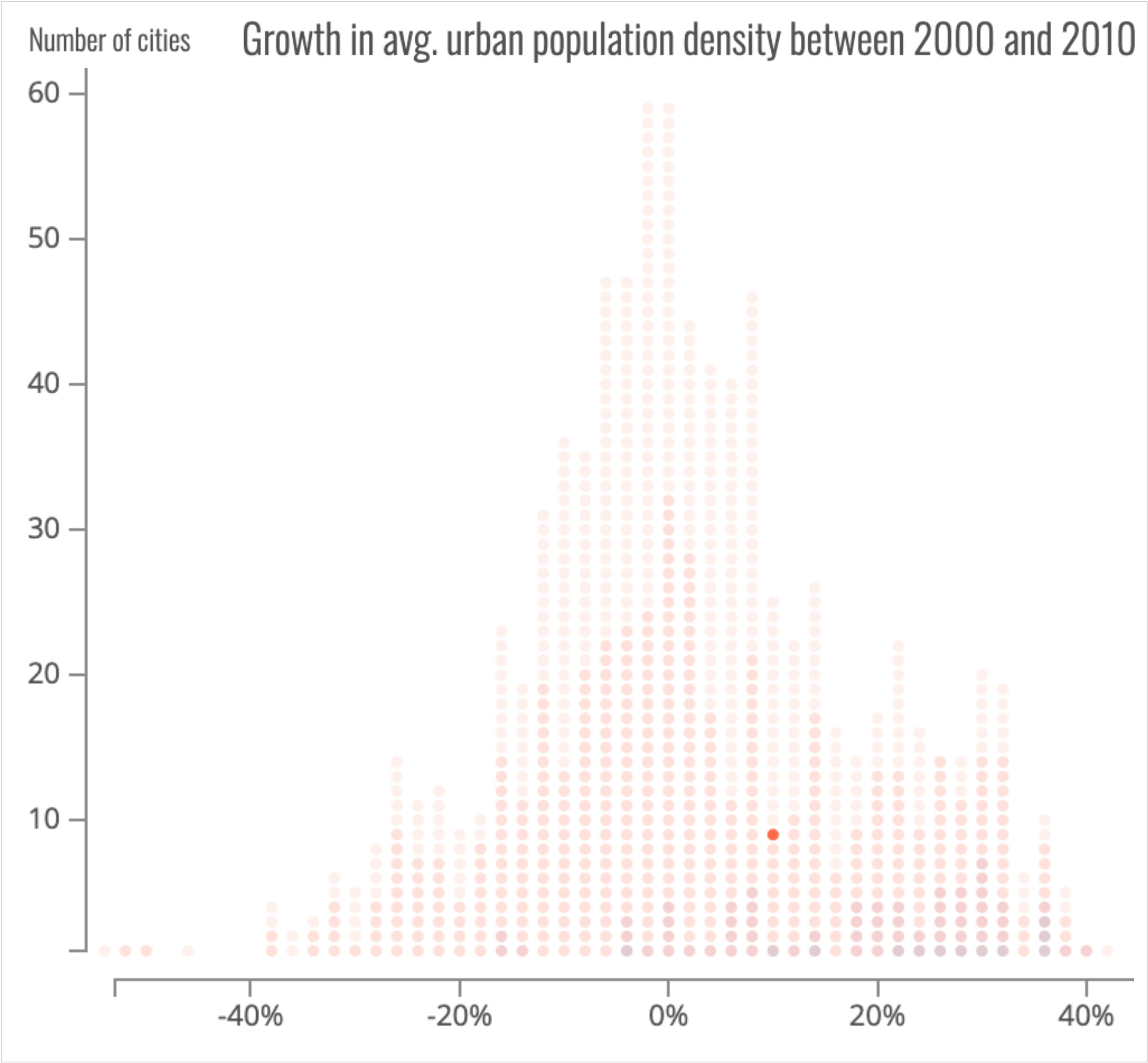
User formulates hypotheses and tests them through various interactions

Demo: moving average

```
let nthPad n xs =
      nth (min (max n 0) (length xs - 1)) xs;
    movingAvg ys window =
      [ sum [ nthPad n ys | n ← [ i - window .. i + window ] ] / (1 + 2 * window)
      | i ← [ 0 .. length ys - 1 ] ];
    movingAvg' rs window =
      zipWith
         (fun x y → {x: x, y: y})
         (map (fun r → r.x) rs)
         (movingAvg (map (fun r → r.y) rs) window);
let points =
      [ { x: r.year, y: r.emissions } | r ← methane, r.type == "Agriculture" ]
in LineChart {
   tickLabels: { x: Rotated, y: Default },
   size: { width: 330, height: 285 },
   caption: "SSP5-8.5 projected methane emissions (Agriculture)",
   plots: [ LinePlot { name: "Moving average", points: movingAvg' points 1 },
            LinePlot { name: "Original curve", points: points } ]
}
```

Next steps

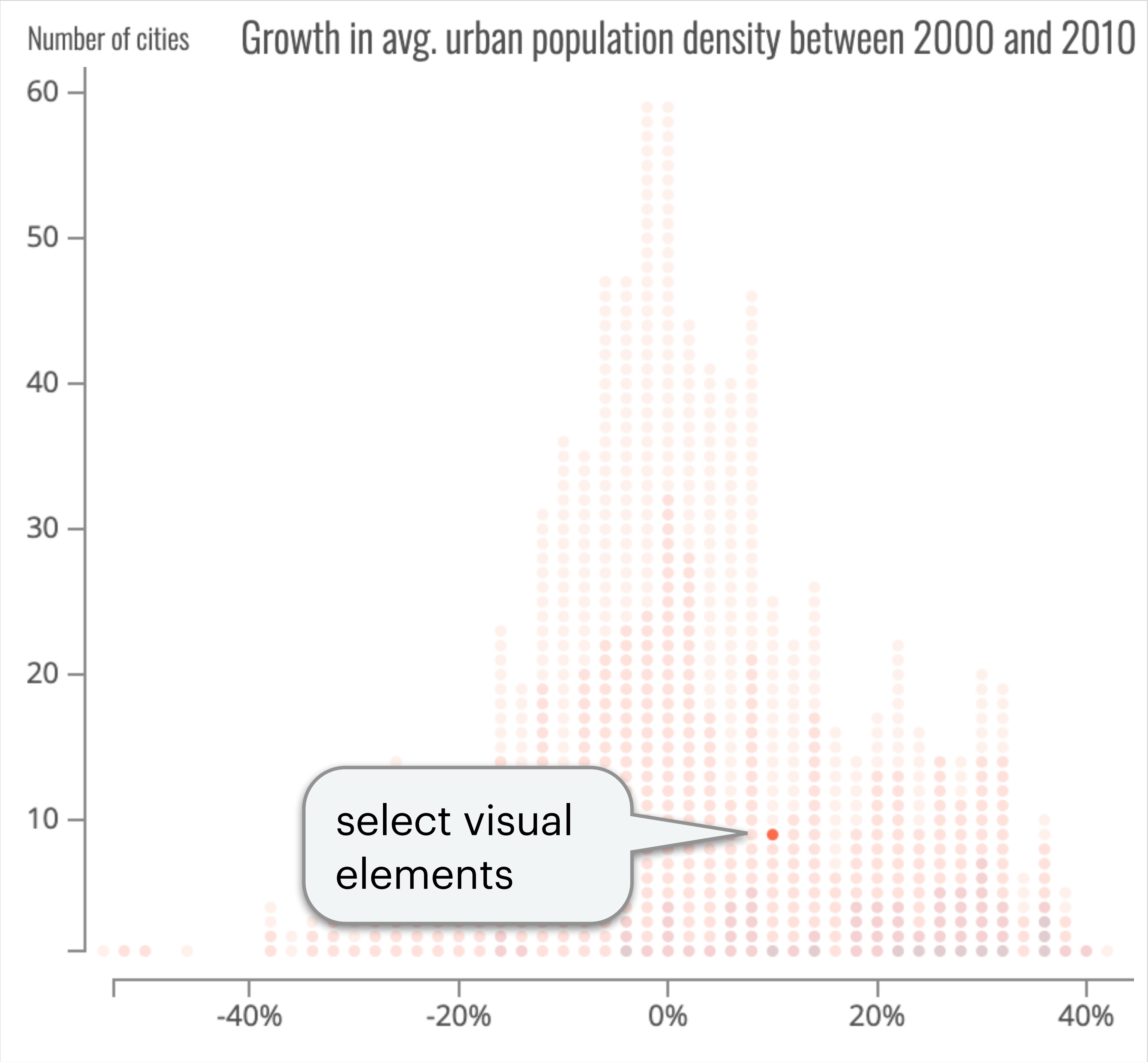
Enrich outputs with
computational explanations
(how, not just what)



Urbanization in East Asia between 2000 and 2010
Nadieh Bremer, Marlieke Ranzijn (http://nbremer.github.io/urbanization, 2015)

Next steps

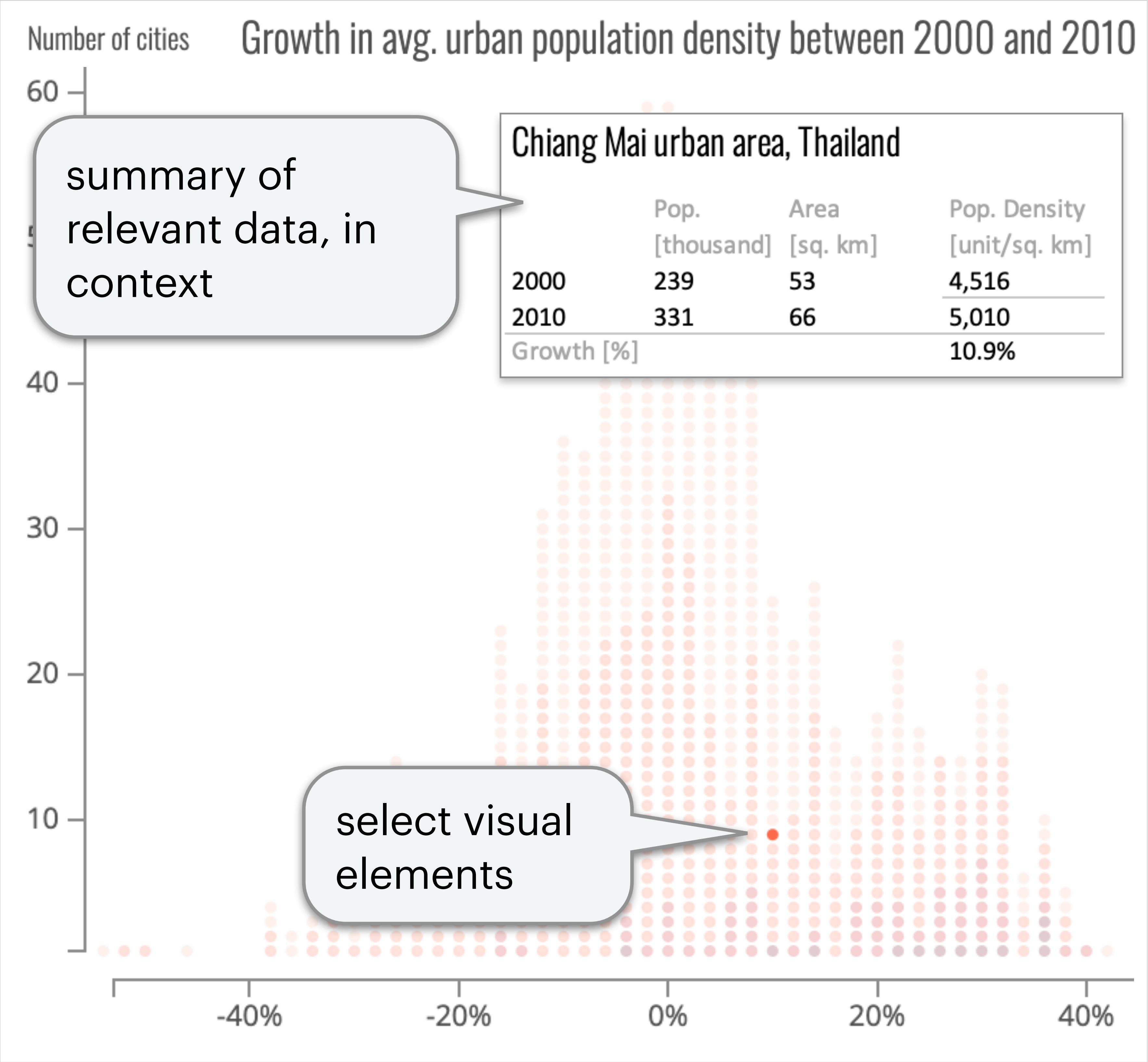
Enrich outputs with **computational explanations** (how, not just what)



Urbanization in East Asia between 2000 and 2010
Nadieh Bremer, Marlieke Ranzijn (http://nbremer.github.io/urbanization, 2015)

Next steps

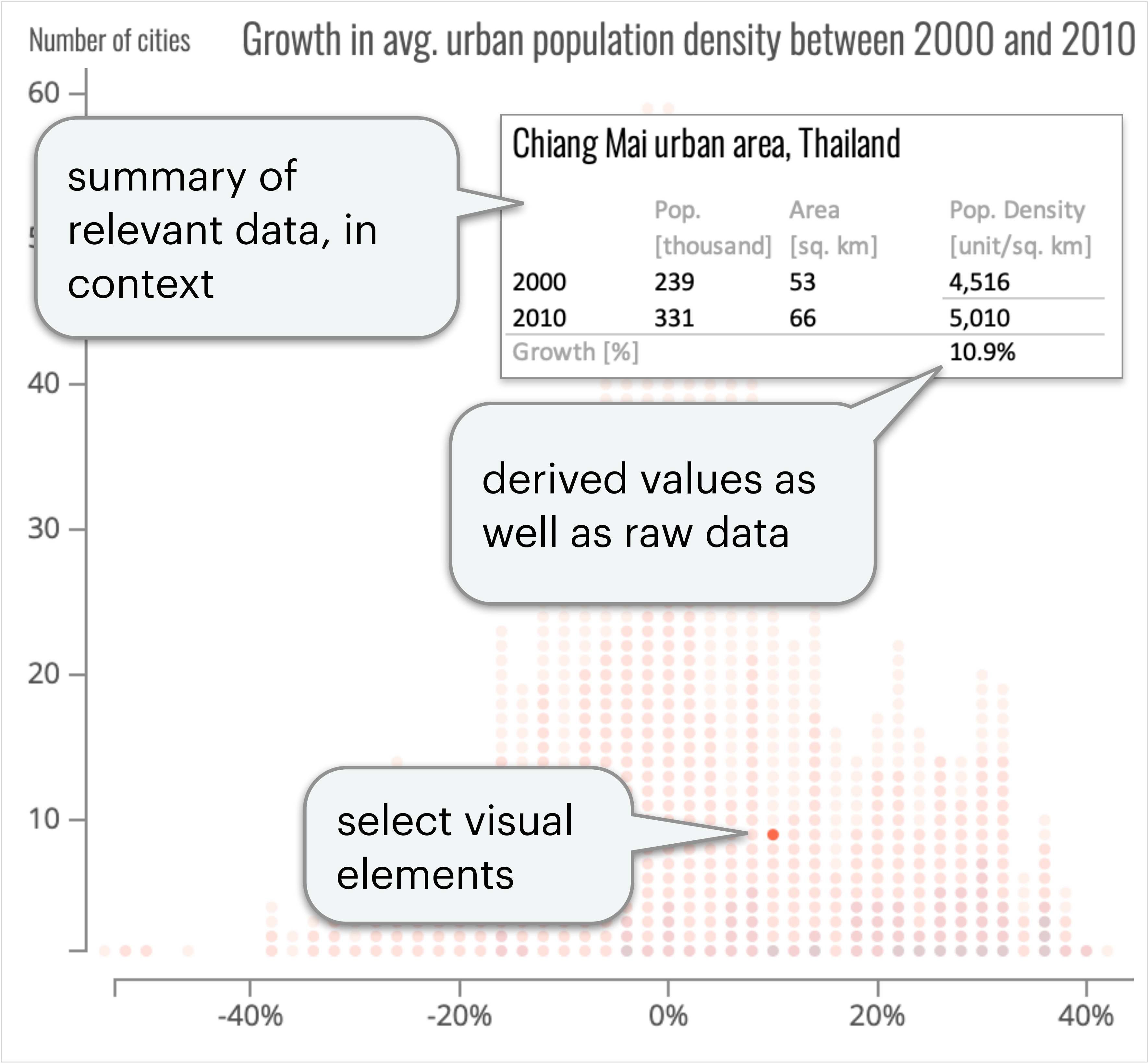
Enrich outputs with **computational explanations** (how, not just what)



Urbanization in East Asia between 2000 and 2010
Nadieh Bremer, Marlieke Ranzijn (http://nbremer.github.io/urbanization, 2015)

Next steps

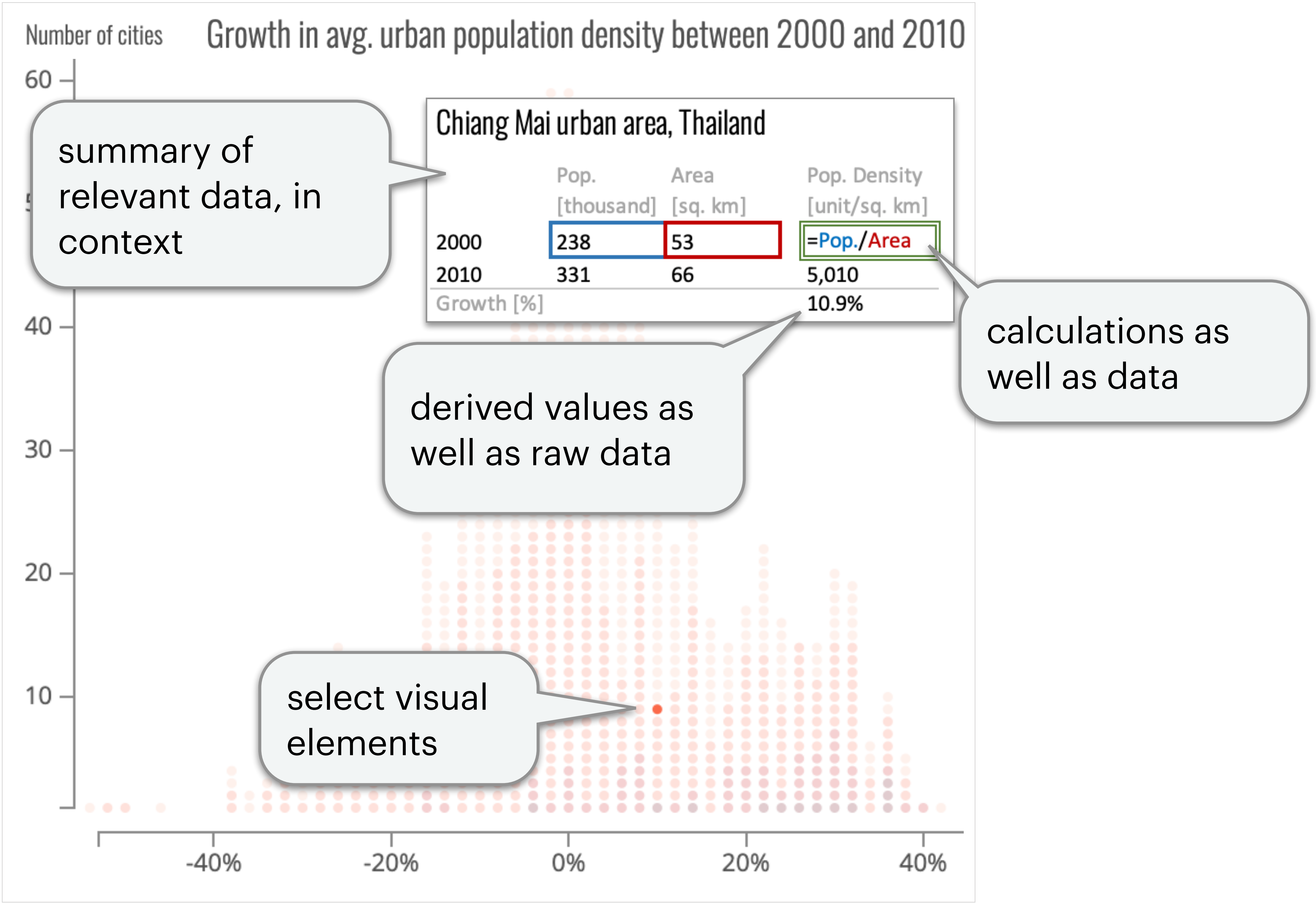
Enrich outputs with **computational explanations** (how, not just what)



Urbanization in East Asia between 2000 and 2010
Nadieh Bremer, Marlieke Ranzijn (http://nbremer.github.io/urbanization, 2015)

Next steps

Enrich outputs with **computational explanations** (how, not just what)

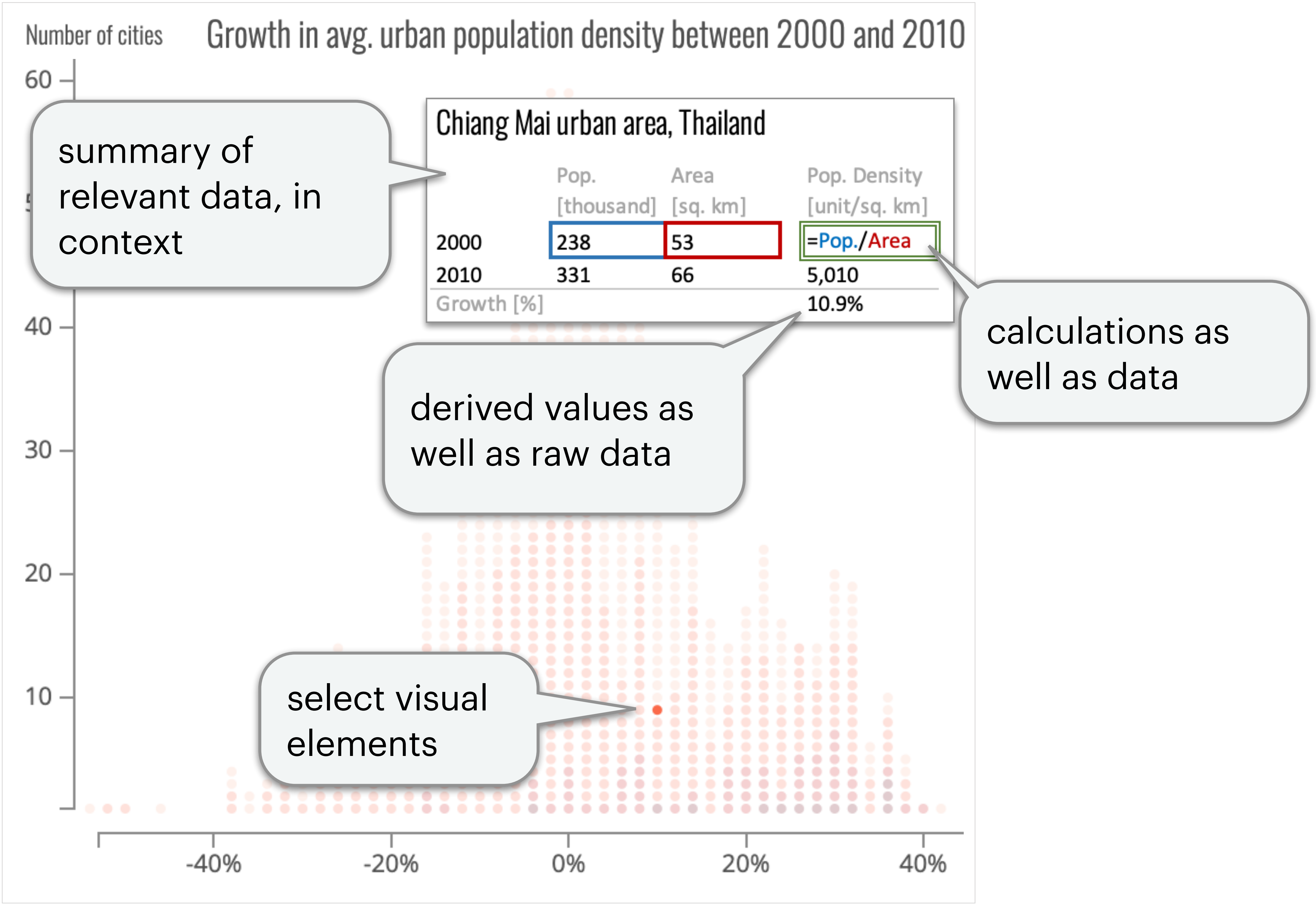


Urbanization in East Asia between 2000 and 2010
Nadieh Bremer, Marlieke Ranzijn (http://nbremer.github.io/urbanization, 2015)

Next steps

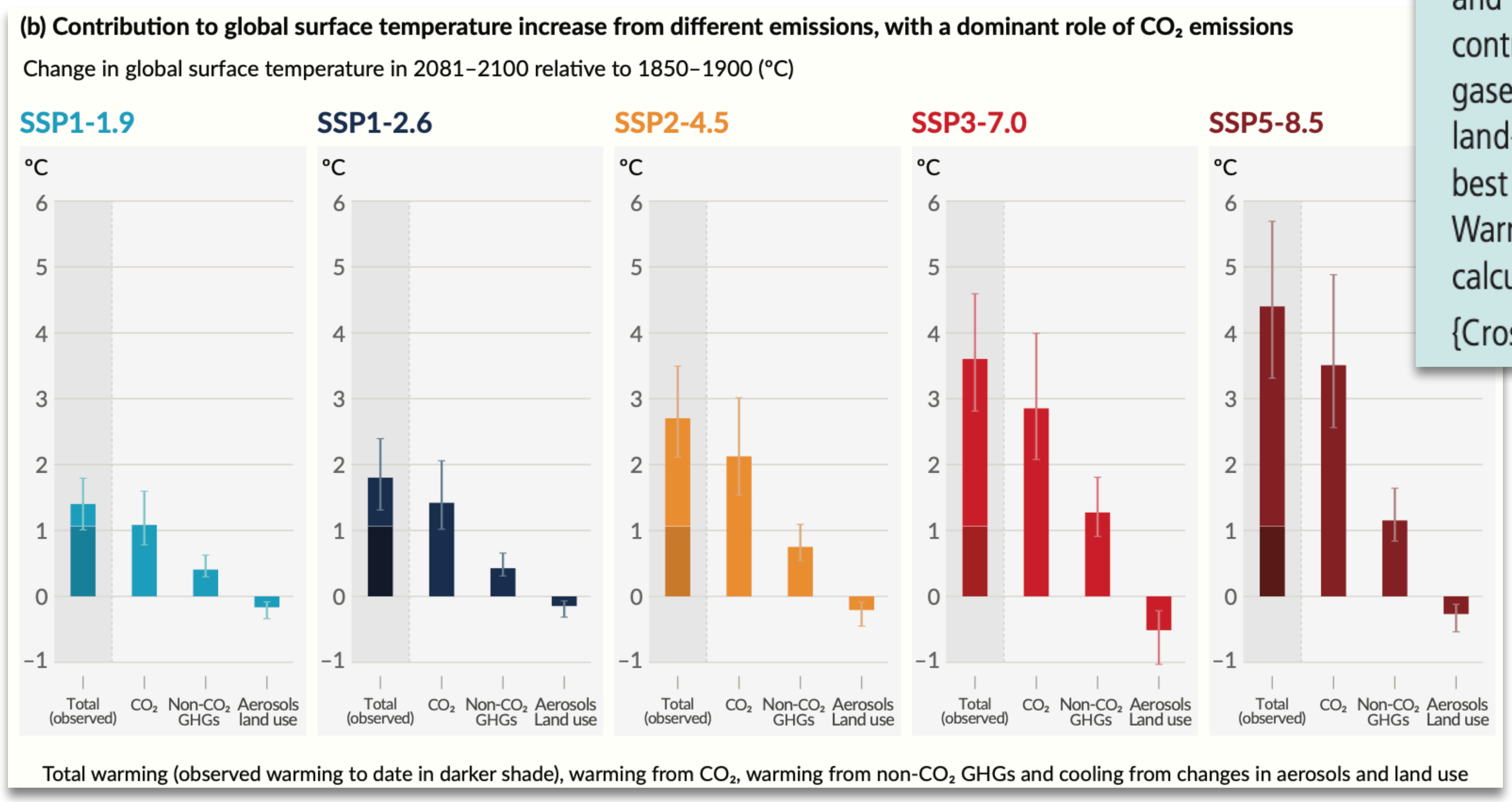
Enrich outputs with **computational explanations** (how, not just what)

cf. moving average example



Urbanization in East Asia between 2000 and 2010
Nadieh Bremer, Marlieke Ranzijn (http://nbremer.github.io/urbanization, 2015)

Next steps

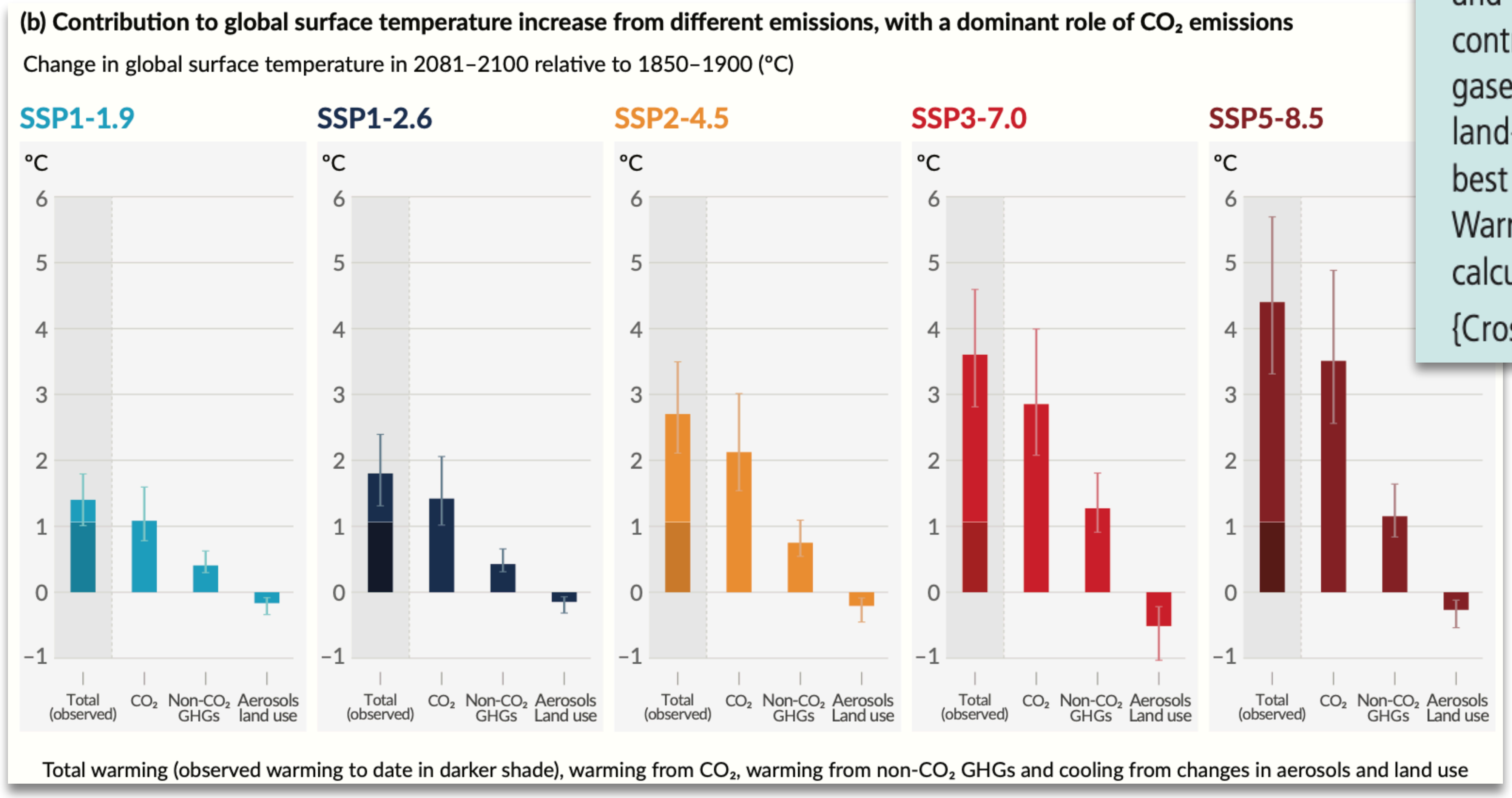


IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Next steps



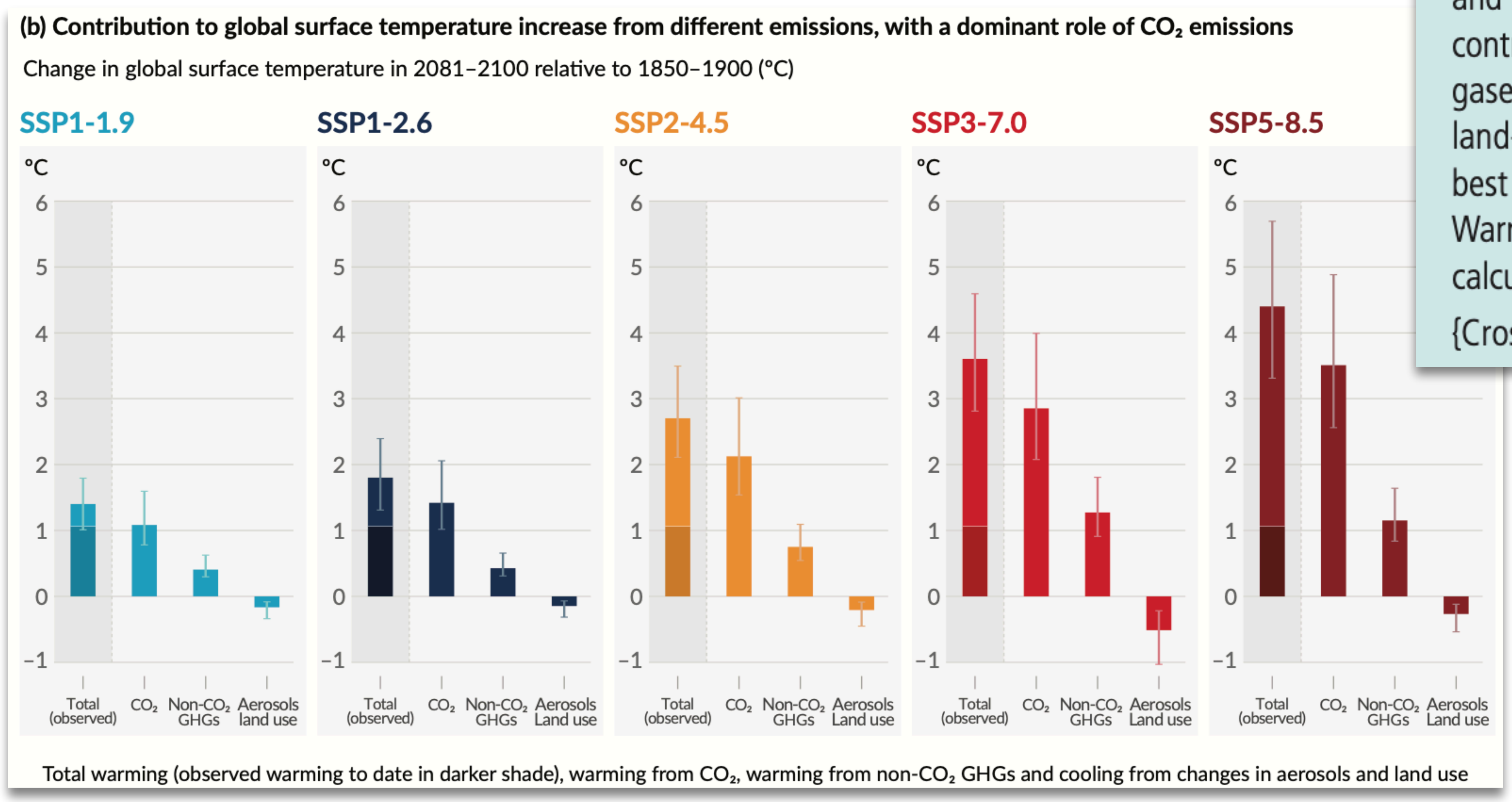
IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

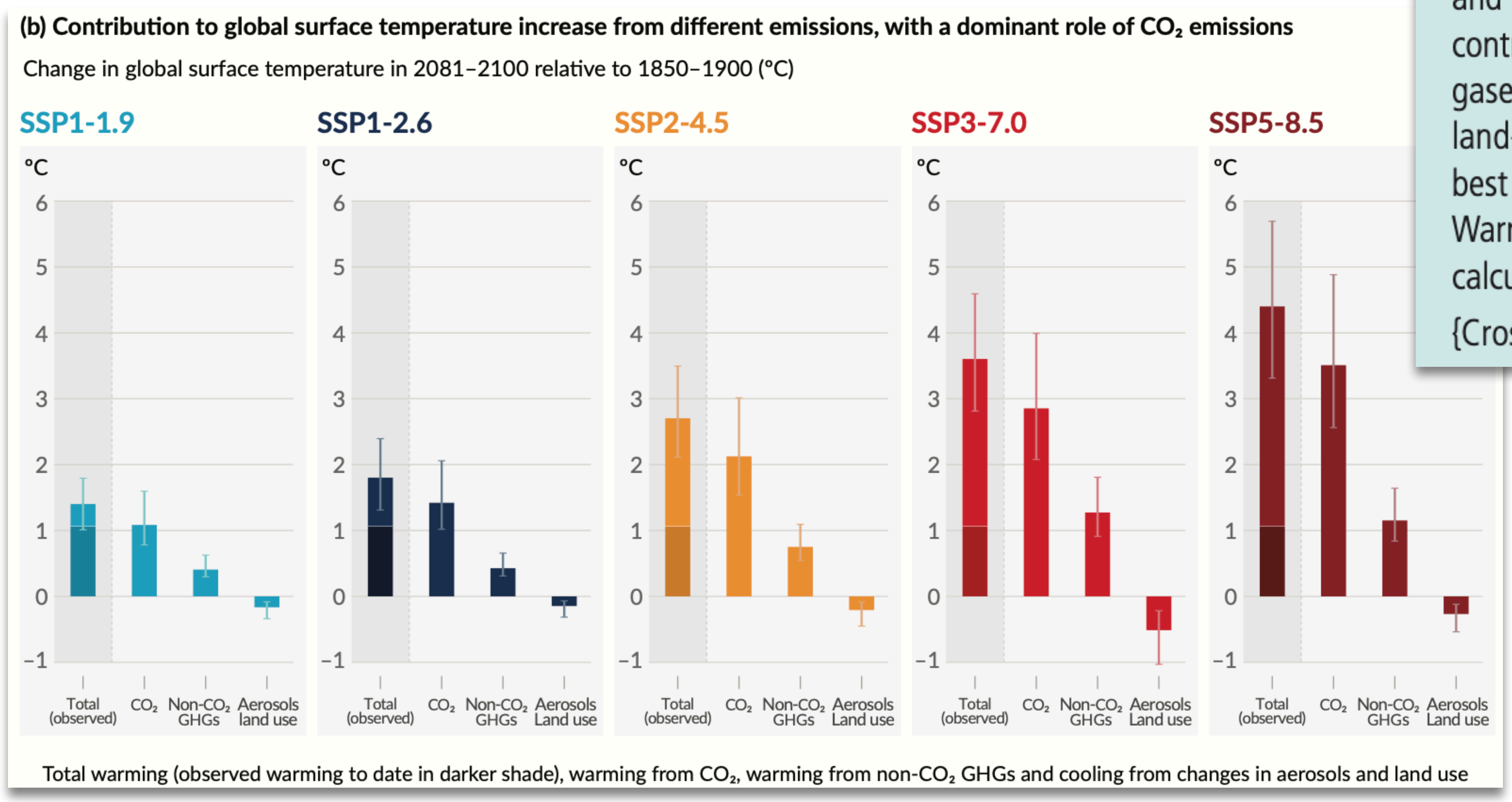
{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Text which is hard to interpret

- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

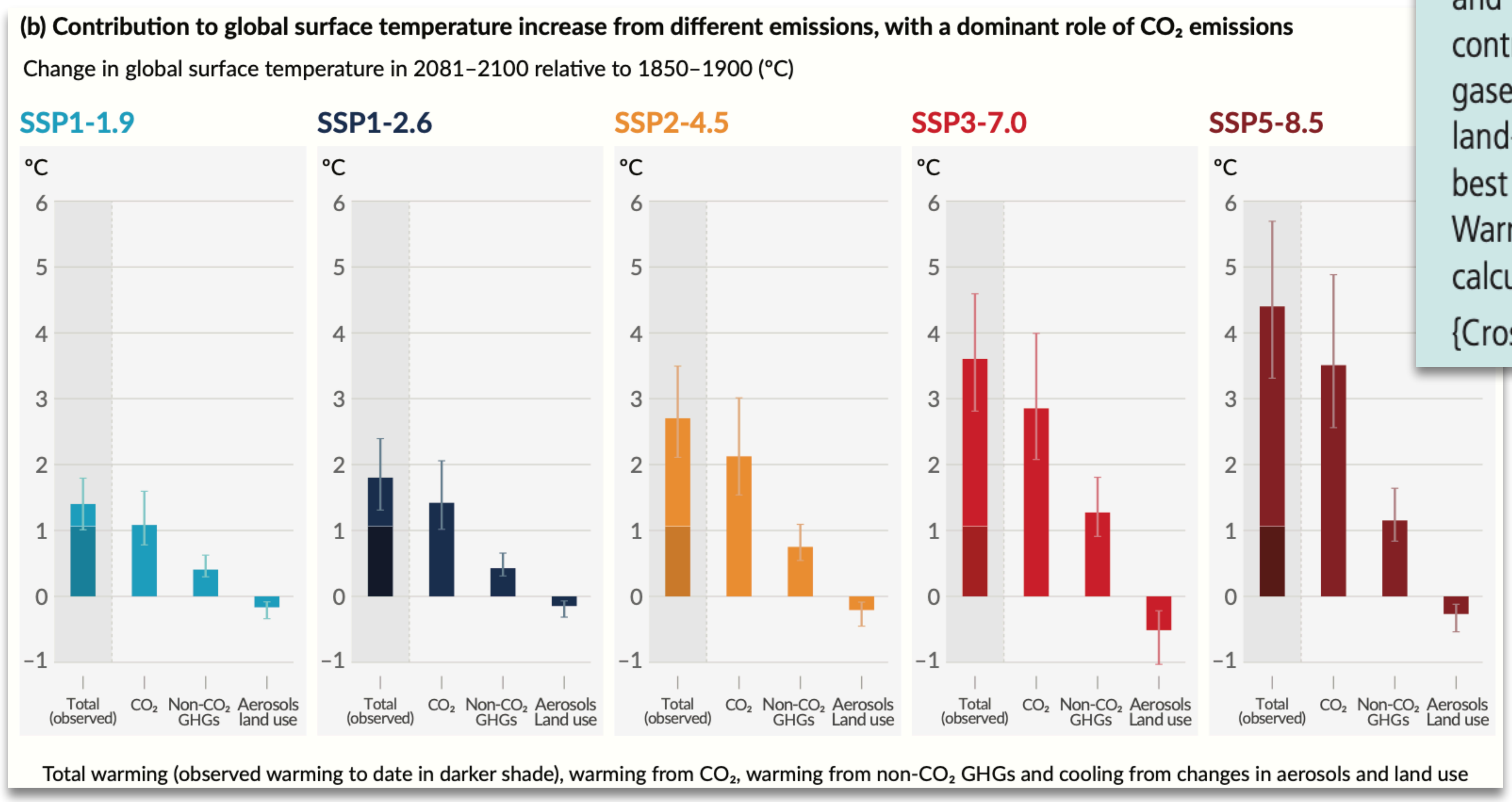
{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Text which is hard to interpret

- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

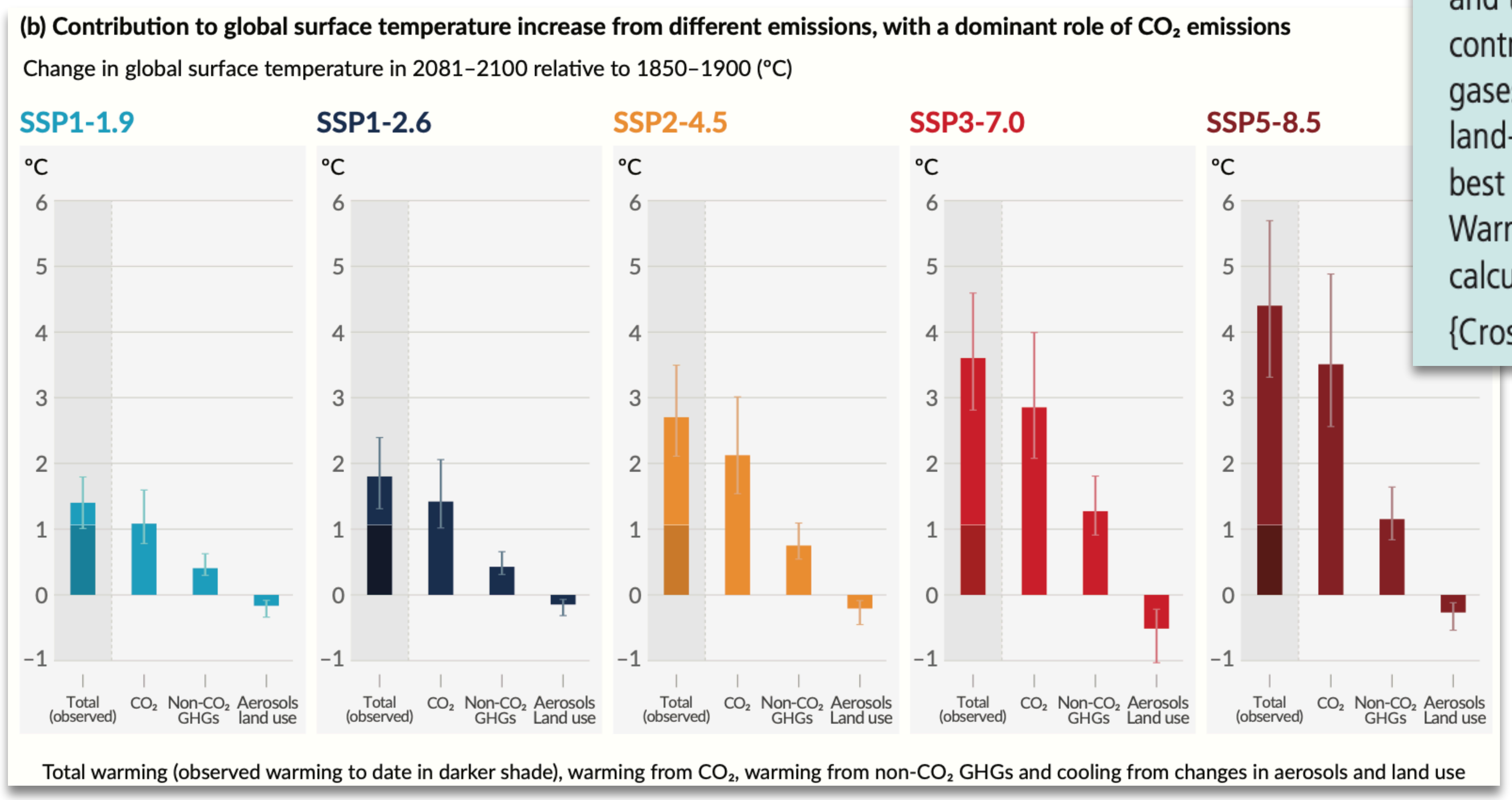
{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Text which is hard to interpret

- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

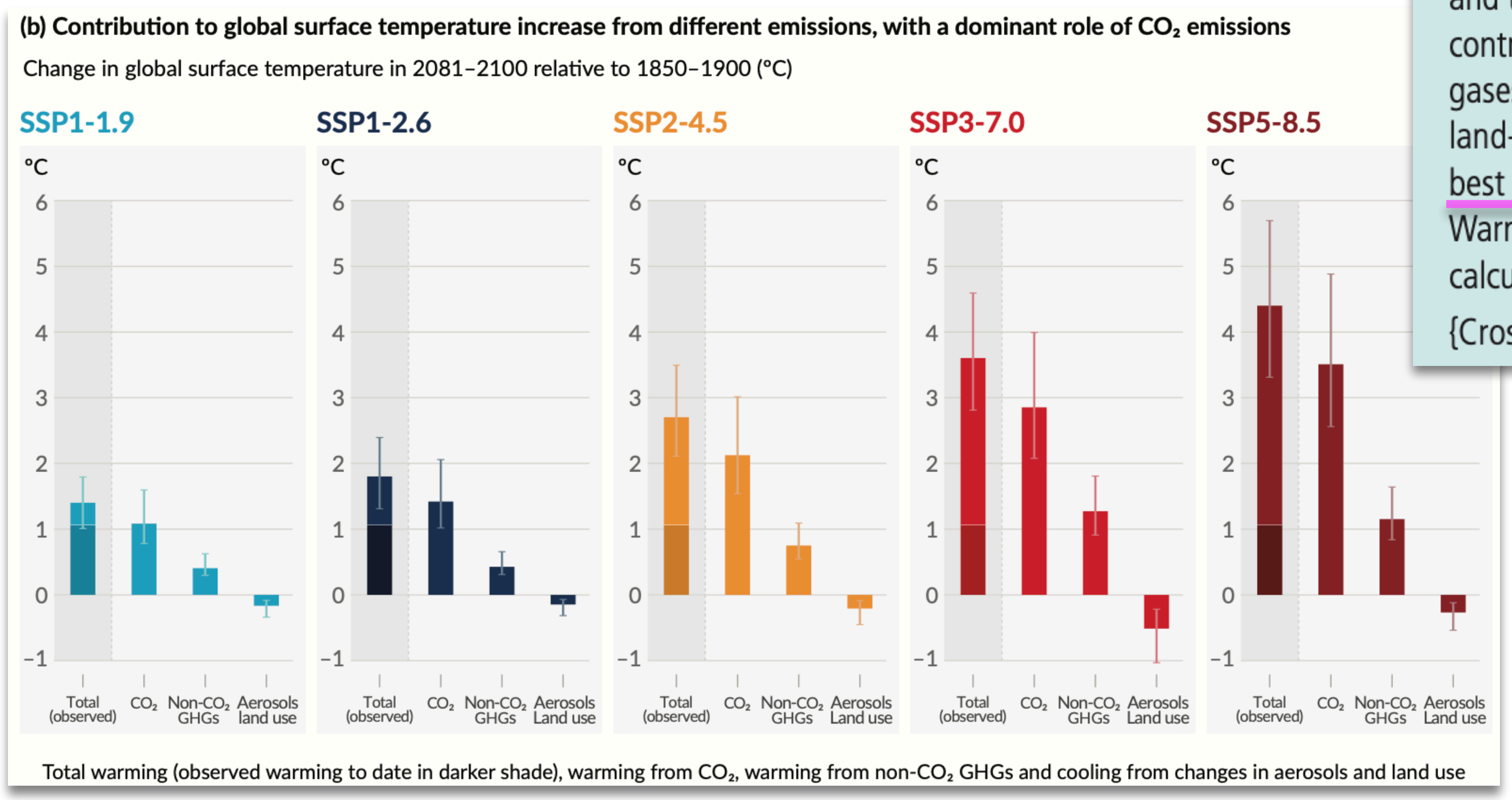
Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Text which is hard to interpret
- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

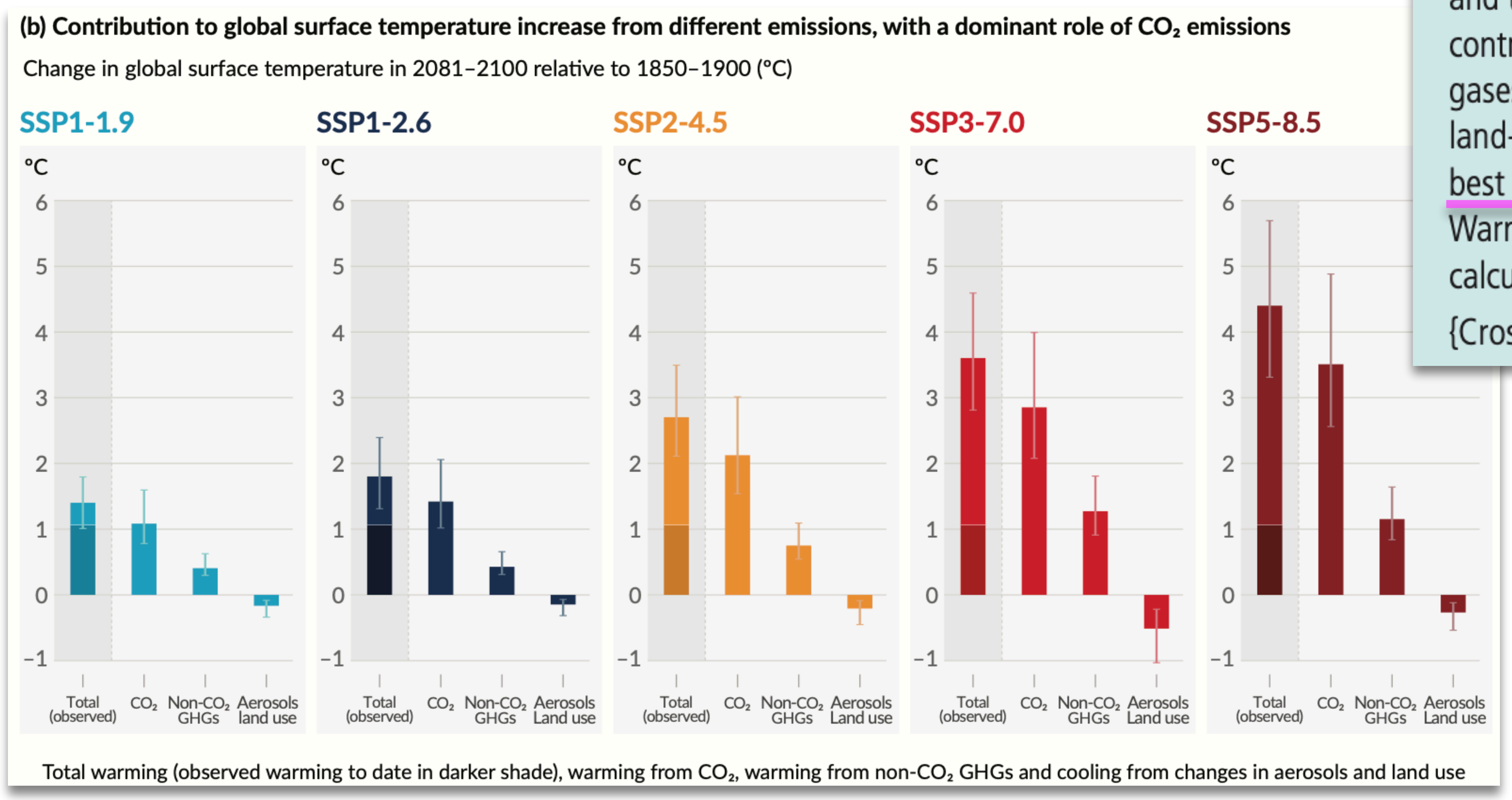
Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Text which is hard to interpret
- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.

{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Research question:

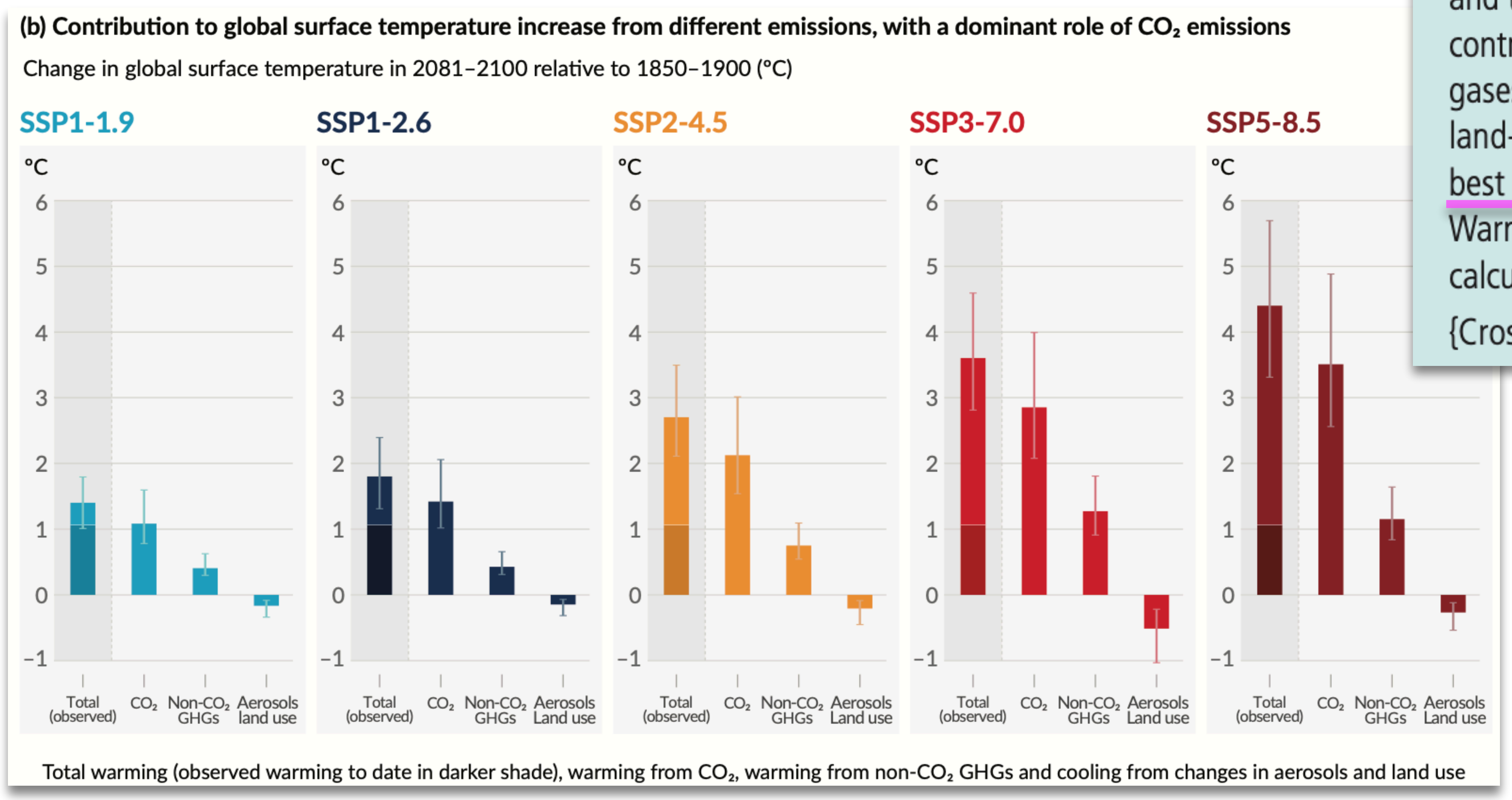
How can we facilitate natural language discourse that is "data-driven"?

- explorable, explainable, verifiable

Text which is hard to interpret

- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Next steps



IPCC Sixth Assessment Report (AR6) WG1, Summary For Policymakers (2021)

Panel (b) Warming contributions by groups of anthropogenic drivers and by scenario are shown as the change in global surface temperature (°C) in 2081–2100 relative to 1850–1900, with indication of the observed warming to date. Bars and whiskers represent median values and the *very likely* range, respectively. Within each scenario bar plot, the bars represent: total global warming (°C; 'total' bar) (see Table SPM.1); warming contributions (°C) from changes in CO_2 ('CO_2' bar) and from non-CO_2 greenhouse gases (GHGs; 'non-CO_2 GHGs' bar: comprising well-mixed greenhouse gases and ozone); and net cooling from other anthropogenic drivers ('aerosols and land use' bar: anthropogenic aerosols, changes in reflectance due to land-use and irrigation changes, and contrails from aviation) (see Figure SPM.2, panel c, for the warming contributions to date for individual drivers). The best estimate for observed warming in 2010–2019 relative to 1850–1900 (see Figure SPM.2, panel a) is indicated in the darker column in the 'total' bar. Warming contributions in panel (b) are calculated as explained in Table SPM.1 for the total bar. For the other bars, the contribution by groups of drivers is calculated with a physical climate emulator of global surface temperature that relies on climate sensitivity and radiative forcing assessments.
{Cross-Chapter Box 1.4; 4.6; Figure 4.35; 6.7; Figures 6.18, 6.22 and 6.24; 7.3; Cross-Chapter Box 7.1; Figure 7.7; Box TS.7; Figures TS.4 and TS.15}

Policy reports, scientific papers and news articles make important claims using **text**

Text which is hard to interpret
- graded adjectives
- iteration
- mereleogy (whole-part)
- quantitative expressions

Research question:
How can we facilitate natural language discourse that is "data-driven"?
- explorable, explainable, verifiable

Working hypothesis (2 key ingredients):
- transparent programming languages
- generative AI for authoring text programmatically

Thank you!

Collaborators

Joe Bond [1]
Haofei Chen [4]
Colin Crawford [4]
Cristina David [1]
Thomas Frith [2]
Harleen Gulati [1]
Hana Iza Kim [2]
Minh Nguyen [1]
Dominic Orchard [5, 2]
Roly Perera [2, 1]
Tomas Petricek [3]
Achintya Rao [6, 2]
Meng Wang [1]

[1]University of Bristol
[2]University of Cambridge
[3]Charles University
[4]University of Edinburgh
[5]University of Kent
[6]University of West of England

https://f.luid.org
https://github.com/explorable-viz/fluid

University of BRISTOL
The Alan Turing Institute
UNIVERSITY OF CAMBRIDGE
Institute of Computing for Climate Science