

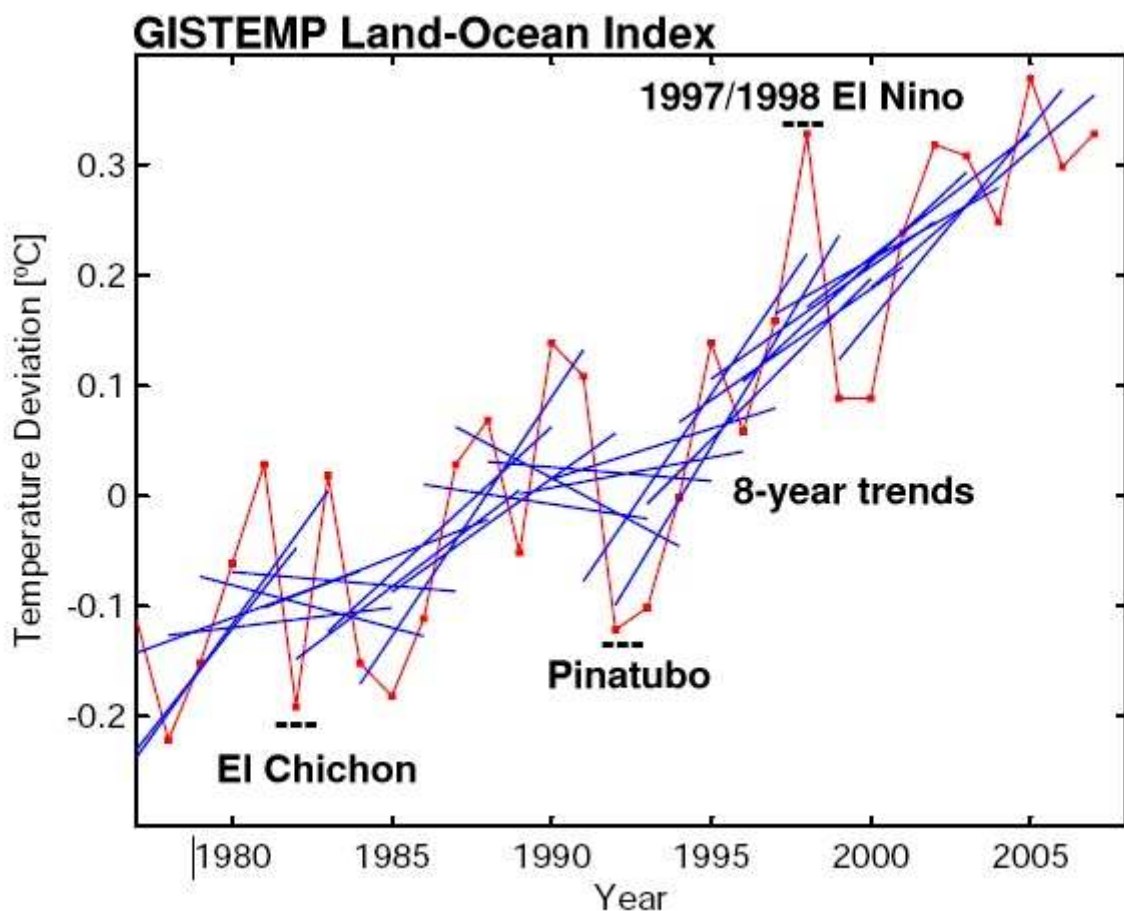
Incertidumbre, ruido y el arte de comparar datos y modelos.

Gavin Schmidt and Stefan Rahmstorf

Traducido por Mario Cuellar para Globalizate (22/01/07)

John Tierney y Roger Pielke Jr, han discutido (1) recientemente los intentos de validar (o falsificar) las proyecciones del IPPC del cambio de la temperatura global durante el periodo 2000-2007. Otros han intentado mostrar que las cifras del último año implican que "El calentamiento global se ha detenido" (2) o que se "ha tomado un respiro" (Uli, Die Wlet) (3). Sin embargo, como la mayoría de nuestros lectores observarán, estas comparaciones son incorrectas porque básicamente comparan el cambio climático a largo plazo con la variabilidad de la meteorología a corto plazo.

Esta aparece claro en cuanto echas un vistazo a la siguiente gráfica:



La línea roja es la temperatura media global anual registrada GISTEMP (4) (aunque con cualquier otro conjunto de datos lo haría también), mientras que las líneas azules son las líneas de tendencia de 8 años – una para cada periodo de 8 años de datos en la gráfica. Lo que muestra es exactamente lo que cualquiera debería esperar: Las tendencias sobre

periodos tan cortos son variables, a veces pequeña, a veces grande, a veces negativa – dependiendo del año con el que comiences. La media de todas las tendencias de 8 años está cercana a la tendencia a largo plazo ($0.19^{\circ}\text{C}/\text{Década}$), pero la desviación estándar es casi tan grande ($0.17^{\circ}\text{C}/\text{Década}$), implicando que una tendencia tendría que ser o $>0.5^{\circ}\text{C}/\text{Década}$ o muchos mas negativa ($<-0.2^{\circ}\text{C}/\text{Década}$) para caer por ello obviamente fuera de la distribución. En consecuencia, comparar tendencias cortas tiene muy poco valor de distinción entre expectativas alternativas. Así, debe quedar claro que comparaciones a corto plazo son erróneas, pero las razones de porque y lo que en su lugar debería hacerse, son valiosas cuando se explican.

Lo primero que hay que apuntar (y realmente es lo que siempre hacemos) es que el sistema climático tiene cantidades enormes de variabilidad en periodos de día a día, mes a mes, año a año y década a década. Gran parte de esta variabilidad (una vez que cuentas el ciclo diario y las estaciones) es aparentemente caótica y sin relación a algún factor externo – es el tiempo. Algunos aspectos del tiempo atmosférico son predecibles – la localización de tormentas de medias latitudes con unos pocos días de adelanto, la progresión del evento El Niño unos pocos meses antes, etc, pero la predictabilidad rápidamente se evapora debido a la extrema sensibilidad de la meteorología a incertidumbres inevitables en las condiciones iniciales. De esta manera, por la mayoría de intentos y propósitos, se puede pensar en el componente meteorológico como aleatorio.

Si estás interesado en el componente forzado del clima –y mucha gente lo está- entonces necesitas establecer el tamaño de la señal forzada esperada en relación al ‘ruido’ meteorológico no forzado. Sin esto, el significado del cualquier cambio observado es imposible de determinar. La señal a la tasa de ruido es realmente muy sensible a al registro climático exacto (o métrica) que estás observando, y si una señal puede ser claramente vista variará enormemente a través de diferentes aspectos del clima.

Un ejemplo obvio es observar la anomalía de la temperatura en una única estación meteorológica. La desviación estándar en la ciudad de Nueva York para la anomalía media mensual es de alrededor de 2.5°C , para la media anual es de 0.6°C , mientras que anomalía de la media global es de 0.2°C . Así, cuanto más largo es el periodo de tiempo de promedios y más amplio es el promedio espacial, más pequeño es el ruido meteorológico y mayor oportunidad de detectar cualquier señal particular.

En el mundo real, hay otras fuentes de incertidumbre que añaden una parte de ‘ruido’ a esta discusión. En primer lugar, existe la incertidumbre de que cualquier métrica particular climática está realmente representando lo que dice ser. Esto puede ser debido a un conjunto de

datos aislados o puede relacionar el procedimiento por el que los datos sin procesar se ponen juntos. Pueden ser aleatorios o sistemáticos y hay un par de buenos ejemplos de esto en las variaciones de los registros de la temperatura en la superficie o cerca de ella.

Errores en el conjunto de datos son fáciles de ver en la diferencia entre el producto de datos GISTEMP de la temperatura superficial (que extrapola sobre la región Ártica) y el producto HADCRUT3v que asume que las anomalías de la temperatura Árticas no se extienden pasada la tierra. Ambas son opciones defendibles, pero cuando calculas las anomalías medias mensuales en una situación donde el Ártico se está calentando rápidamente, hay una obvia eliminación entre los dos registros (y realmente GISTEMP ha tenido una tendencia mayor). Sin embargo, las tendencias a largo plazo son muy similares.

Un error más sistemático es observar las diferencias entre las versiones RSS y UAH del registro de la temperatura observada por satélites MSU-LT (en la baja troposfera). Ambos grupos están nominalmente intentando estimar lo mismo de los mismos datos, pero a causa de las asunciones y los métodos usados para atar todo lo que los diferentes satélites mostraban, puede haber grandes diferencias en las tendencias. Dado que solo tenemos dos ejemplos de esta métrica, la incertidumbre sistemática verdadera es claramente más grande que la simple diferencia entre ellas.

Lo que estamos persiguiendo es como evaluar nuestra comprensión de que está dirigiendo el cambio climático como encapsulado en los modelos del sistema climático. Aquellos modelos pensados pueden ser tan simples como una tendencia extrapolada, o tan complejas como un el más avanzado de la tecnología actual, el GCM. Cualquiera que sea la fuente de una estimación de lo que 'debería' estar pasando, hay tres asuntos que necesitan ser apuntados:

- Primero, ¿están las implementaciones cambiando como esperábamos? Es muy fácil predecir que un peatón será probablemente golpeado si se pone en el camino de un camión, pero la predicción solo puede ser validada si realmente se pone en medio! En el caso del clima, necesitamos saber como de bien estimamos los forzamientos (gases invernadero, efectos volcánicos, aerosoles, radiación solar, etc) en las proyecciones.
- Segundo, ¿Qué es la incertidumbre en la predicción dada para un forzamiento particular? Por ejemplo, ¿Cuántas veces se salva nuestro pobre peatón porque el camión pudo esquivarle? Para los cambios de temperatura esto es equivalente a la incertidumbre en las tendencias proyectadas a largo plazo. Esta incertidumbre depende de la sensibilidad climática, la longitud del tiempo y el tamaño de la variabilidad no forzada.

- En tercer lugar, necesitamos comparar como con como y tener cuidado sobre que cuestiones estamos realmente preguntando. Esto se ha vuelto fácil con el archivo (5) de la simulación de modelos para el siglo XX (comentaremos algo mas en el futuro).

Es muy valioso extenderse sobre el tercer punto ya que a menudo es el que más quebrados de cabeza provoca a la gente. En las proyecciones de un modelo, es ahora una práctica estándar hacer varias simulaciones que tengan diferentes condiciones iniciales para extender el rango de estados posibles del tiempo. Cualquier simulación individual tendrá el mismo forzado cambio climático, pero tendrá un resultado diferente del ruido no forzado. Al promediar sobre varias ejecuciones, el ruido (que no está correlacionado en cada una) se elimina, y lo que queda es una estimación de la señal forzada y su incertidumbre. Esto es de alguna manera, análogo al promedio de todas las tendencias de corto plazo en la figura de arriba, y como allí, puedes a menudo obtener una muy buena estimación del cambio forzado (o la media a largo plazo),

Aunque los problemas pueden ocurrir si la estimación del cambio forzado es comparada directamente a la tendencia real para ver si son consistentes. Necesitas recordar que el mundo real consiste en una tendencia forzada (potencial) pero también un componente meteorológico aleatorio. Esto fue un asunto reciente del documento Douglass et al (6), donde decían que las observaciones estaban fuera de la media de la tendencia troposférica del modelo y su incertidumbre. Confundieron la incertidumbre, en como de bien podemos estimar la señal forzada (la media de todos los modelos) con la distribución de tendencias+ruido.

Esto podría parecer confuso, pero la analogía del lanzamiento del dado podría ser útil. Si tienes un puñado de dados normales ('modelos') entonces el valor de la media es 3.5 con una desviación estándar de ~ 1.7 . Por tanto, la media sobre 100 lanzamientos tendrá una distribución de 3.5 ± 0.17 lo que quiere decir que es una estimación bastante buena. Para asegurar si otro dado está trucado no es suficiente solo comparar con una tirada de ese dado. Por ejemplo, si obtienes un 5, que está significativamente fuera del valor esperado derivado de los 100 lanzamientos anteriores, pero está claramente dentro de la distribución esperada.

Llevándolo a los modelos climáticos, puede haber un fuerte acuerdo de que $0.2^\circ\text{C}/\text{Década}$ es el valor esperado de la actual tendencia forzada, pero comparando con la tendencia real simplemente a ese número más o menos la incertidumbre en su valor es incorrecto. Esto es lo que está implícitamente siendo hecho en la cifra del post de Tierney.

Si esta no es la mejor forma de hacerlo, ¿Cuál es el mejor camino? Bien, si comienzas a coger tendencias más largas, entonces la incertidumbre

en la tendencia estimada se aproxima a la incertidumbre en la tendencia esperada, punto en el cual comienza a tener sentido compararlas ya que el componente "meteorológico" ha quedado fuera de los promedios. En el registro de la temperatura global superficial, esto ocurre para tendencias mayores que 15 años, pero para áreas más pequeñas con nivel más altos de ruido (como la Antártida), el periodo de tiempo puede ser de varias décadas.

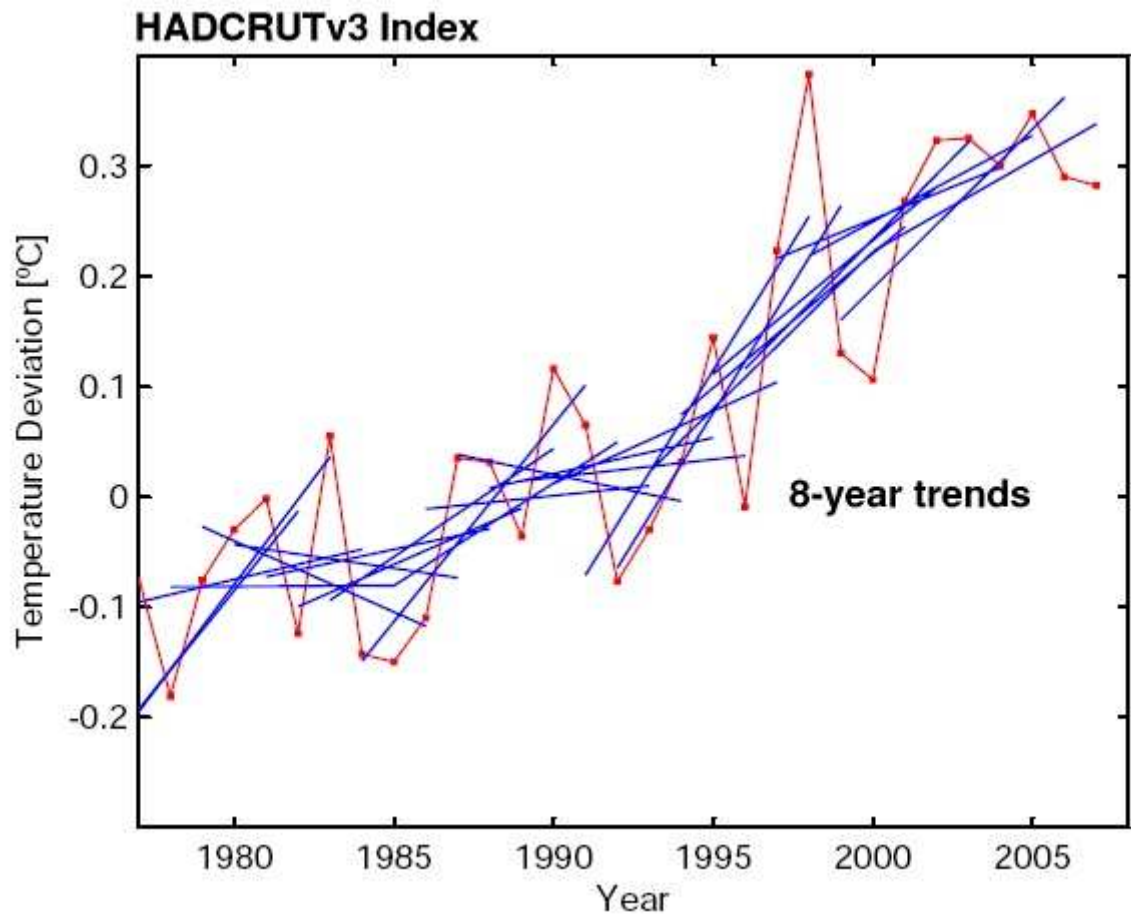
¿Está la gente volviendo a las primeras proyecciones y comprobando lo buenas que son? Si. Hemos hecho eso para las proyecciones de Hansen (7) en 1988, Stefan y colegas lo hicieron para el CO2, las proyecciones de temperatura y el nivel del mar en el IPCC TAR (Rahmstorf et al, 2007) (8), y el IPCC lo hizo aplicándolo en la figura 1.1 de AR4 Capítulo 1 (9). Cada uno de estos análisis muestran que cuanto el plazo es más largo, las tendencias de temperatura son realmente lo que se espera. El aumento del nivel mar, por otro lado, parece estar subestimado por los modelos pero las razones aún no están claras.

Finalmente, este asunto parece que ha salido de la expectativa de que algún evento meteorológico a corto plazo durante los próximos años será definitivamente la prueba de que el calentamiento global antropogénico es un problema o no. Como la discusión anterior debería dejar claro no es la cuestión correcta a preguntar. En su lugar, la cuestión debería ser, ¿Se harán análisis en los próximos años que mejorarán la evaluación de los modelos climáticos? Aquí la respuesta probablemente sea sí. Habrá mejores estimaciones de las tendencias a largo plazo de la precipitación, nubosidad, vientos, intensidad de las tormentas, profundidad del hielo, retroceso de glaciares, calentamiento del océano, etc. Tenemos expectativas de lo que aquellas tendencias deberían ser, pero en muchos casos el 'ruido' es todavía muy grande para aquellas métricas sean una ligadura útil. Cuando pasa el tiempo, el ruido en las tendencias largas disminuye y lo que aparece entonces determinará lo bien que comprendemos lo que está pasando.

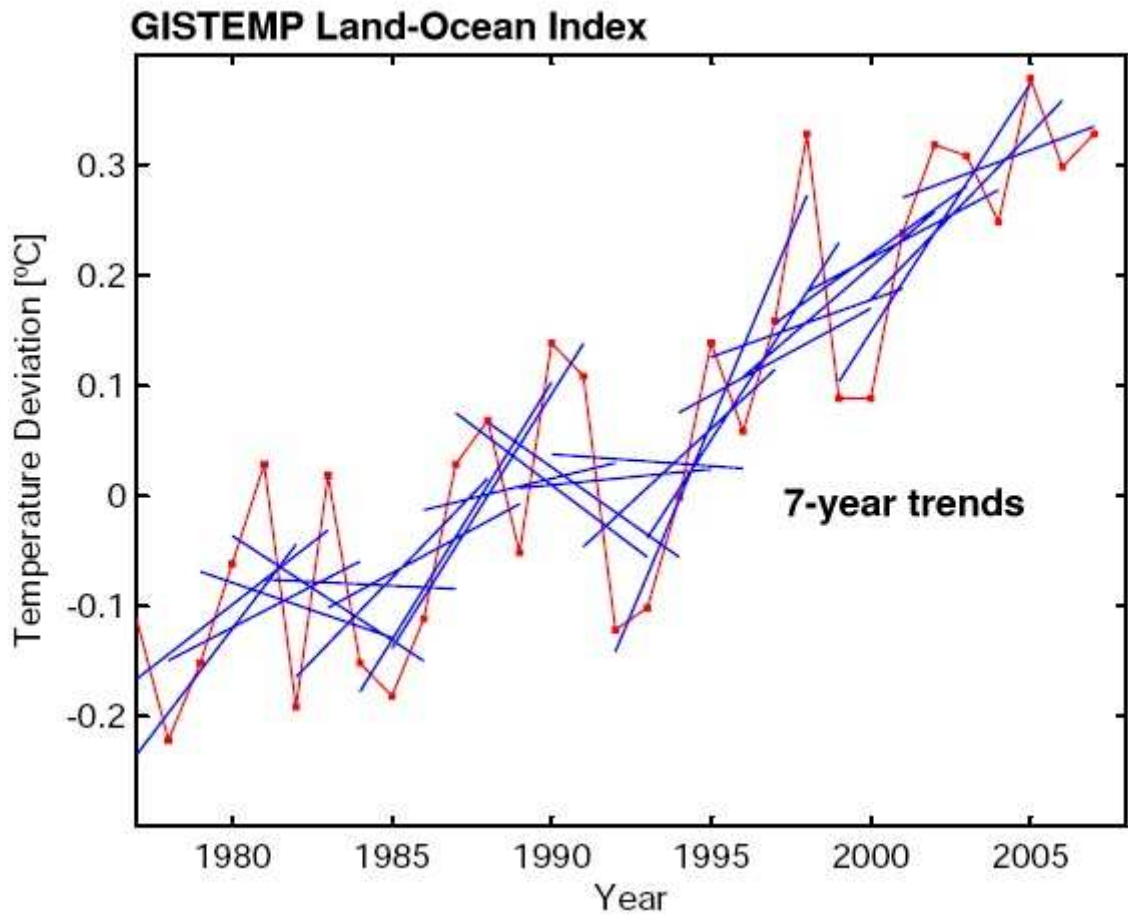
Actualización.

Nos complace ver que ha habido gran interés en este post. Varios lectores nos han pedido gráficos adicionales. Aquí están:

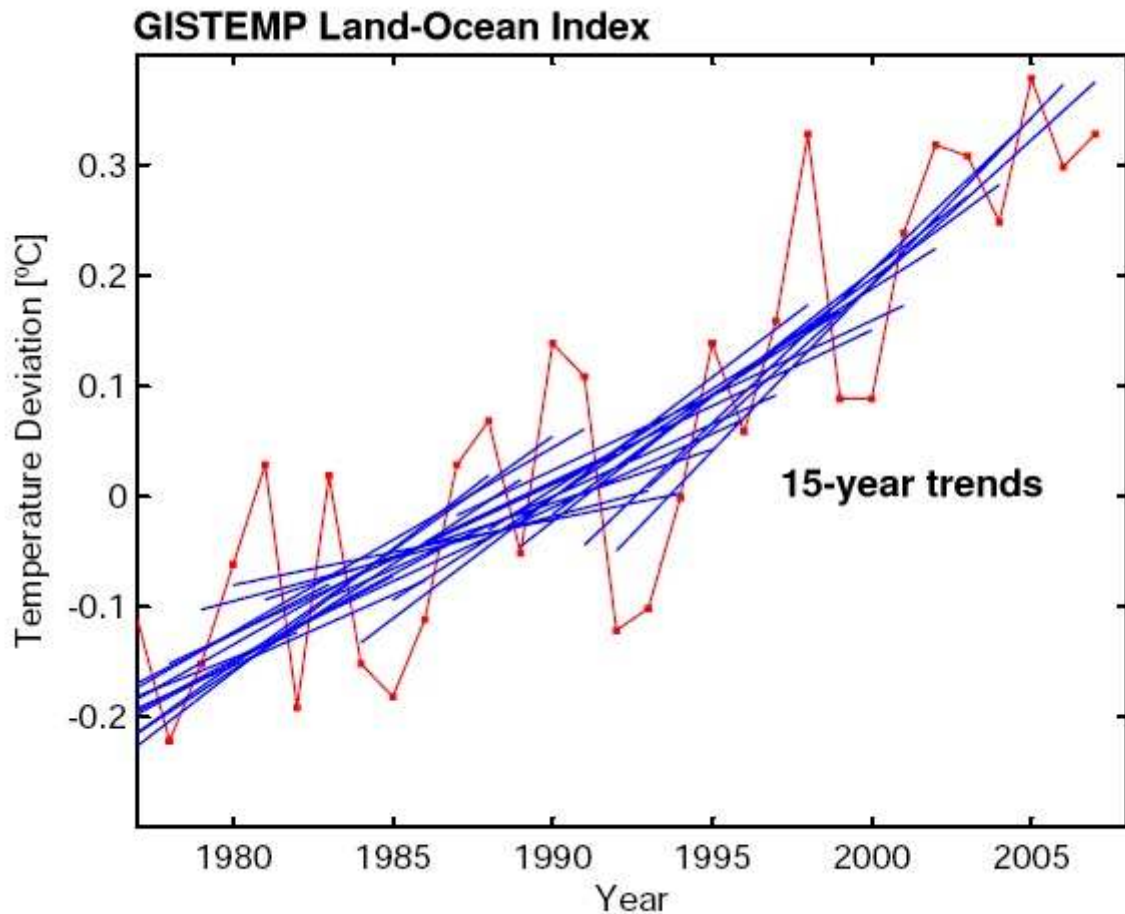
- UK Met Office data (En lugar de los datos GISS) con líneas de tendencia de 8 años



- GISS data con líneas de tendencia de 7 años (en lugar de 8 años)



- GISS data con líneas de tendencia de 15 años



Estas gráficas ilustran que las tendencias de 8 años en los datos del Met Office Inglés son, por supuesto, justo tan ruidosas como los datos GISS; que las líneas de tendencia de 7 años son, claro está, incluso más ruidosas que las de la tendencia de 8 años; y que las cosas comienzan a estabilizarse (tendencias que obtienes robustez estadística) cuando se usa el promedio de 15 años. Esto ilustra el punto claro que estábamos intentando hacer; observar solo 8 años de datos es mirar primariamente al "ruido" de la variabilidad interanual más que a la tendencia a largo plazo forzada. Esto tiene tanto sentido como analizar las observaciones de temperatura del 10 al 17 de abril para chequear si realmente hay calentamiento durante la primavera.

Hay una última actualización (10) de la comparación de la temperatura global con las proyecciones del IPCC TAR (Tahmstorf et al., Science 2007) con los valores de 2007 añadidos (11). Con ambos conjuntos de datos las tendencias a largo plazo observadas todavía están por encima en la parte superior del rango proyectado por el IPCC.

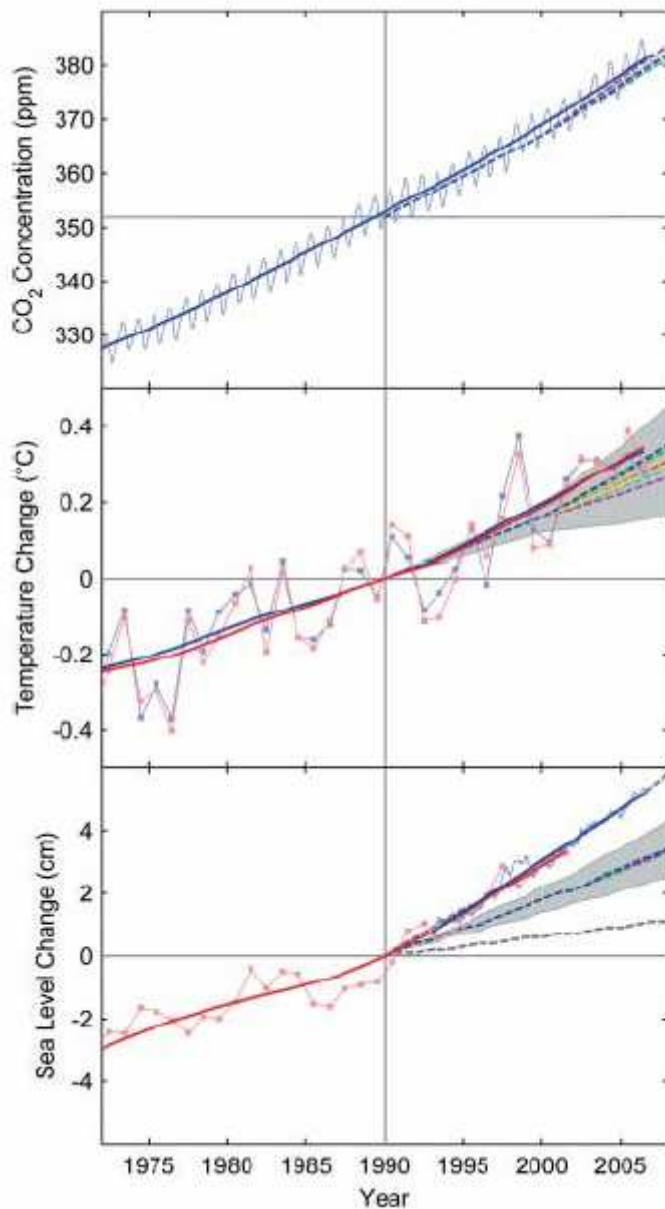


Figura. Cambios en los principales parámetros climáticos globales desde 1973 comparados con los escenarios del IPCC (mostrados como líneas discontinuas y rangos grises. **Arriba:** Concentración de dióxido de carbono mensual y su línea de tendencia en Mauna Loa, Hawai (azul) hasta enero de 2007 de Scripps en colaboración con NOAA, (unidades en partes por millón, ppm). **Medio:** Temperatura superficial global media anual del océano y la tierra combinada del GISS (roja) y del Hadley Center Climate Research Unit (azul) hasta 2006, con tendencias. **Bajo:** Datos de nivel del mar basados principalmente en dispositivos de mareas (anual, rojo) y desde altímetros de satélites (datos de 3 meses, azul hasta mediados de 2006).

Sobre los autores:

Stefan Rahmstorf

Físico y oceanógrafo. Ha dirigido el capítulo paleoclimático en el cuarto informe del IPCC.

Gavin A. Schmidt

Climatólogo del Instituto Goddard de la NASA.

Artículo original:

http://www.realclimate.org/index.php/archives/2008/01/uncertainty-noise-and-the-art-of-model-data-comparison/langswitch_lang/sp#more-523

REFERENCIAS

(1) <http://tierneylab.blogs.nytimes.com/2008/01/10/a-spot-check-of-global-warming/>

(2) <http://www.newstatesman.com/200712190004>

(3)

http://www.welt.de/wissenschaft/article1453165/Die_Erderwaermung_macht_erst_mal_Pause.html

(4) <http://data.giss.nasa.gov/gistemp/>

(5) http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php

(6) <http://www.realclimate.org/index.php/archives/2007/12/tropical-troposphere-trends/>

(7) <http://www.realclimate.org/index.php/archives/2007/05/hansens-1988-projections/>

(8)

http://www.pikpotsdam.de/~stefan/Publications/Nature/rahmstorf_etal_science_2007.pdf

(9) http://ipcc-wg1.ucar.edu/wg1/Report/AR4WG1_Print_Ch01.pdf

(10)

<http://www.pikpotsdam.de/~stefan/material/globaltemperatures.pdf>

(11)

http://www.pikpotsdam.de/~stefan/Publications/Nature/rahmstorf_etal_science_2007.pdf