

Artículo de Divulgación Científica

Desarrollo y uso de sistemas numéricos de pronóstico estacional

Introducción

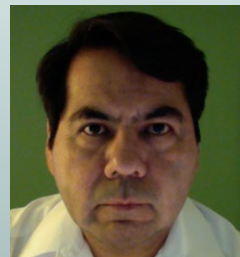
Los sistemas numéricos de pronóstico climático producen las guías numéricas y gráficos que se utilizan en los centros de predicción estacional. Los sistemas procesan datos a través de una cadena de algoritmos que va desde el control de calidad de las observaciones, extracción de información observacional, hasta el pronóstico y despliegue de mapas representando posibles estados futuros del clima. El valor que tienen estos productos depende del usuario pero para los pronosticadores este depende tanto de la calidad y contenido de la información del pronóstico como de su confiabilidad y qué tan oportunamente el producto llega a sus manos. El pronóstico en sí no tiene un valor intrínseco a menos que se le utilice para tomar decisiones ante un evento anticipado (Murphy, 1993).

Al menos cuatro capacidades tecnológicas deben existir para tener un pronóstico climático exitoso: una red de observaciones capaz de muestrear eficazmente los procesos relevantes en el clima, un esquema numérico que acopie y resuma los datos presentes y pasados y los convierta en una serie consecutiva de campos tridimensionales de variables climáticas consistentes entre sí según las leyes de la física, un método o modelo que proyecte el estado presente a un estado futuro y finalmente un procedimiento estadístico que convierta las salidas del modelo a un producto para el usuario. Estas capacidades han avanzado aceleradamente en las últimas décadas e integradas a un sistema de predicción, se les considera una de las revoluciones científicas más importantes de la historia (Bauer et al., 2016), si bien existen desafíos científicos, tecnológicos y económicos para continuar avanzando y haciéndolas más cercanas a las necesidades de los usuarios.

La ciencia y la tecnología del pronóstico del tiempo y el clima continua expandiéndose tanto en lo referente al entendimiento de los principales fenómenos que dan lugar al estado promedio y variabilidad del clima como en lo referente a los nuevos sistemas de observación, procesamiento de datos, y generación de productos derivados. A continuación se describen brevemente algunos conceptos esenciales que permiten reconocer la ciencia, métodos y procedimientos que subyacen en la predicción estacional desde la perspectiva numérica operativa.

Diferencia entre sistema y modelo de pronóstico

Los sistemas numéricos de pronóstico del tiempo y estacional están compuestos de una serie de procedimientos lógicos, representados en algoritmos que extraen la información contenida en los datos de las redes de observaciones, típicamente dispersas en el tiempo y espacio, la filtran y la transforman en representaciones de las variables atmosféricas y/o oceánicas en mallas regulares, produciendo lo que se conoce como “análisis”. Estos análisis son



Dr. Malaquías Peña
Centro de Modelación Ambiental del NWS/
NOAA(EE.UU.)

Dr. en Meteorología de la Universidad de Oklahoma y Maryland, EE.UU. Es profesor asociado en el departamento de Ingeniería Civil y del Medio ambiente en la Universidad de Connecticut (EE.UU); es ingeniero consultor en el Centro de Modelación Ambiental del NWS/NOAA. Se graduó como Físico en la Universidad Nacional Autónoma de México (UNAM). Es autor de artículos de alto impacto citados por más de 2000 publicaciones de acuerdo con Google Scholar. Su investigación se centra en las áreas de la modelación numérica y predicción del clima. Ha sido consultor a través de la OMM para la promoción de recursos humanos y capacidades de pronóstico numérico en meteorología en Latinoamérica.

subsecuentemente transformados al espacio matemático del modelo de pronóstico para luego ser proyectadas a futuro.

Estos pronósticos son después transformados de regreso al espacio de las observaciones donde es interpolado y ajustado estadísticamente hasta generar el producto final requerido. En contraste con un sistema de pronóstico, un modelo de pronóstico es la síntesis del conocimiento de la evolución del clima que se obtiene al adaptar las ecuaciones que gobiernan la dinámica de la atmósfera, el océano y otros sub-sistemas, a las estadísticas de observaciones de análisis históricos; su propósito es llevar un estado del clima, representado en su espacio matemático, en un tiempo dado a otro estado futuro del mismo. Dos aspectos son importantes de considerar aquí. Uno es que el problema de pronóstico en un modelo es en esencia un problema de valor inicial (o problema de Cauchy) que es resuelto para cada elemento de la malla. El segundo aspecto es que las salidas del modelo están en el espacio del modelo y regresarlas al espacio físico (por ej. corregir los errores sistemáticos del modelo) es un problema que generalmente involucra relaciones empíricas y herramientas estadísticas tal como el Teorema de Bayes. Por otro lado, el objetivo de un sistema de pronóstico es el de extraer toda la información disponible, procesarla hasta tener un producto final en el menor tiempo posible, y realizar los múltiples procesos eficientes y confiables.

Este último punto es lo que diferencia entre un sistema rutinario o experimental y un sistema operativo. En un sistema de pronóstico operativo se busca tener una confiabilidad del 100% de que el producto estará listo a la hora convenida. Esto implica fuertes demandas en los sistemas de cómputo tanto para correr los procesos a altas velocidades como para tener suficiente redundancia y esquemas de contingencia en caso de que haya problemas durante la integración del modelo.

Elementos típicos de un sistema de pronóstico

En la figura 1, se muestra un diagrama de las componentes típicas de un sistema de pronóstico separado en tres bloques. Estos procesos numéricos se repiten generalmente cada seis horas. A continuación describimos cada proceso.

Desarrollo y uso de sistemas numéricos de pronóstico estacional

Peña, M.

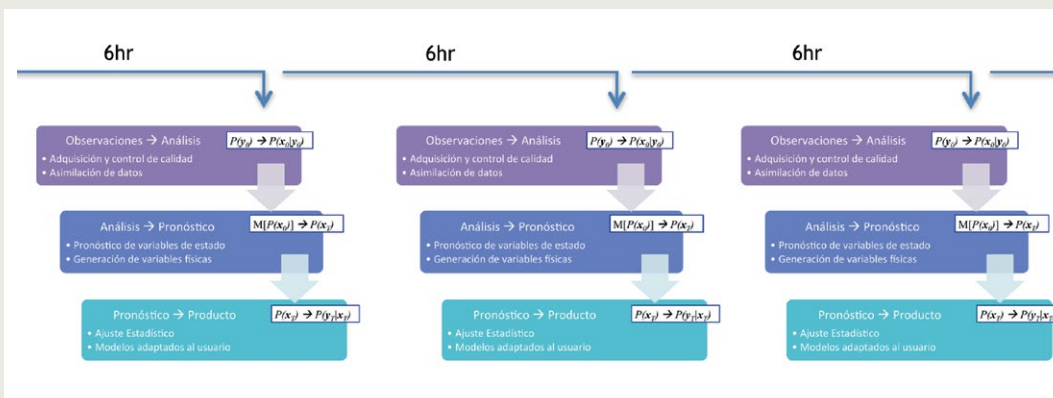


Figura 1. Esquema del proceso de datos en un sistema de pronóstico estacional separado en tres etapas. La primera en observaciones y análisis, que consiste en transformar los datos de las observaciones a una malla regular llamada análisis. La segunda etapa, la cual ocurre en el espacio del modelo y que consiste en propagar el estado del modelo de un tiempo 0 (análisis) a un tiempo T futuro (pronóstico). La tercera etapa consiste en convertir el pronóstico de regreso al espacio físico.

La adquisición y control de calidad de datos, aunque menos glamorosas que otras componentes del sistema tal como la asimilación de datos, son de suma importancia para proveer la información de las redes de observación tanto al esquema de asimilación como para la generación de bases de datos climáticos. Para el caso de la atmósfera y datos de superficie, el NCEP utiliza el sistema MADIS (<https://madis.ncep.noaa.gov>) que decodifica los datos y los codifica en un formato común de unidades y sello de tiempo. En componentes climáticas del modelo tal como el océano y el hielo, la información es procesada de igual manera por diversos grupos encargados de los análisis correspondientes. En tiempo real, los datos son copiados y pasan por una serie de procesos dependiendo del tipo de observación que toma alrededor de 10 minutos.

Los controles de calidad de datos van desde los más simples (nivel 1) en donde se verifica que los valores observados caigan dentro de un intervalo de tolerancia prescrito, a un segundo nivel en donde se verifica que los valores tengan consistencia interna, temporal e histórica-espacial, hasta un tercer nivel en donde se verifica que los valores sean consistentes con los valores de datos vecinos usando la técnica de Interpolación Óptima (Gandin 1963). Dentro de los esquemas de asimilación de datos también existen otros controles de calidad. En estos controles, los valores del dato observado son comparados con un campo preliminar ("first guess") correspondiente al mismo tiempo y región geográfica, y según la diferencia se decide si el dato se toma en cuenta en el análisis o no. Para entender la importancia del control de calidad dentro del esquema de asimilación debe notarse que el esquema genera el análisis como un promedio pesado entre las observaciones y el campo preliminar. Los pesos que utiliza el esquema de asimilación se modelan a partir de estimar la varianza de los errores tanto de las observaciones como del campo preliminar. Si la distribución de errores no coincide con el error real de la observación, es posible que el esquema le dé un peso mayor o menor a lo que debería respecto al peso del campo preliminar, o que el sistema no converja a una solución correcta (Purser, 2011). Esto produciría un pronóstico erróneo.

La piedra angular de los sistemas de pronóstico es el esquema de asimilación de datos, el cual produce los análisis que sirven tanto como condiciones iniciales para el pronóstico, como campos de referencia durante el proceso

de evaluación del pronóstico. Información detallada de los esquemas puede obtenerse en Kalnay (2002). Como se había indicado antes, el esquema de asimilación asigna pesos relativos entre las observaciones y el campo preliminar para calcular el análisis. El promedio pesado dará un análisis determinista. En forma general y considerando las observaciones y el campo preliminar como distribuciones de probabilidad ($P(x)$ y $P(y)$, respectivamente), la asimilación de datos se le puede ver como un problema estadístico de estimación condicionada. En este caso la estimación del estado del clima, representado por el campo preliminar, está condicionado a las observaciones que están disponibles en el tiempo del análisis: $P(x_0|y_0)$, donde el subíndice 0 indica el tiempo del análisis. Los esquemas de asimilación han avanzado considerablemente con el uso de métodos Monte-Carlo (referidos como métodos de ensemble) para la estimación del estado del clima y su incertidumbre. Un área aún en desarrollo es la de asimilación de datos para modelos acoplados océano-atmósfera. Hasta ahora, un esquema es aplicado a cada componente del modelo climático. Así tenemos un análisis para la atmósfera, otro para el océano, etc. Los análisis se acoplan durante la integración del modelo acoplado (e.g., CFSv2, Saha et al., 2014) en el tiempo. A este método se le llama "weakly coupled". Hay altas expectativas en los métodos "strongly coupled", en donde observaciones en el océano afectan el análisis en la atmósfera; o en donde observaciones en la atmósfera afectan al análisis del océano.

Una vez que el campo preliminar corregido o condicionado por las observaciones se transforman al espacio del modelo M, el siguiente paso es integrar a futuro el modelo de pronóstico partiendo de estas condiciones iniciales proporcionadas por el análisis. En este caso $M[x_0] = xT$ para un pronóstico determinista, o $M[P(x_0)] = P(xT)$ denotando la proyección del tiempo 0 al tiempo T para el caso de ensembles. Existen varias clases de modelos de pronóstico, incluyendo modelos analíticos, empíricos y dinámicos, y en cada clase existen diferentes grados de sofisticación o jerarquía en la representación de los procesos físicos observados. Los modelos mas exhaustivos se componen de sub-modelos acoplados entre sí cada uno representando las componentes del sistema natural climático tal como el océano, la atmósfera, la litosfera, etc. A estos modelos se les han dado diferentes nombres como Coupled Global Circulation Models (CGCMs) para diferenciarlos de los modelos estadísticos; actualmente se refieren a ellos como modelos del sistema de la Tierra

(Earth System Models; ESM) porque también incluyen componentes que realizan procesos geoquímicos. La figura 2, muestra las componentes típicas de un ESM y los rangos donde se espera el mayor beneficio.

Continuando con la figura 1, una vez que el pronóstico de las variables de estado ha completado o mientras se está completando, las variables se transforman al espacio físico y se generan cientos de variables a partir de las variables de estado y ecuaciones de diagnóstico. En la última etapa, llamada post-procesamiento, las variables se ajustan estadísticamente y, dependiendo de los usos, estas variables alimentan otros modelos anidados, para pronóstico regional, o “downstream” para el pronóstico hidrológico, etc. Es importante considerar que las salidas de los modelos de pronóstico en el NCEP se colocan de manera automática en los servidores FTP públicos con la expectativa de que los usuarios realicen su propio post-procesamiento y evaluación.

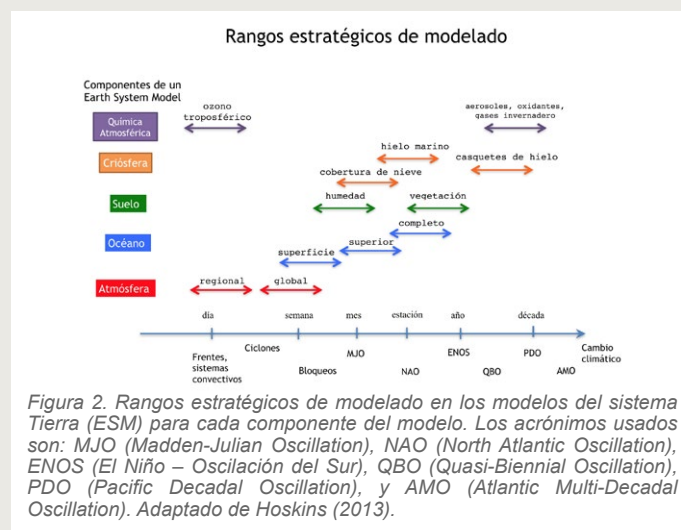
Actualmente también se ponen a disposición tanto los “reanalysis” (análisis de los años pasados) como los reforecasts (o también llamados hindcasts, que son pronósticos retrospectivos), para post-procesar efectivamente los pronósticos crudos.

Fuentes de predicción

El largo y costoso proceso de realizar un pronóstico estacional por medios numéricos debe ser justificado con resultados que demuestren que los pronósticos tengan habilidad predictiva superior a la de los pronósticos basados en métodos empíricos tales como persistencia o promedio histórico. Otro método empírico es el de los análogos. La idea fundamental es que lo que ocurrió en el pasado puede ocurrir en el futuro si se tuvieran condiciones idénticas del estado climático. La dificultad es que es imposible encontrar estados análogos en la limitada serie de datos históricos considerando los billones de diferentes variantes o grados de libertad que definen completamente el estado del clima natural. Tomaría del orden de 1030 años encontrar análogos de los campos típicos en el dominio hemisférico (van den Dool 1994). El hecho de que la atmósfera es caótica, y consecuentemente las demás componentes climáticas acopladas lo son, hace que una pequeña diferencia entre el análisis y su análogo pueda crecer exponencialmente con el tiempo resultando en dos escenarios completamente distintos. En otras palabras, una pequeña diferencia en el estado inicial puede no significar una diferencia pequeña en el estado final. El principio de análogos también se utiliza para hacer los ajustes estadísticos del pronóstico en la etapa de post-procesamiento. Por ejemplo, para hacer correcciones de sesgo, se utilizan promedios con datos históricos y se espera que las estadísticas del pasado se puedan aplicar los pronósticos.

En contraste con los análogos, en el caso de Earth System Models se asume que las leyes físicas que gobiernan al clima no cambian con el tiempo. Las principales fuentes de predictibilidad provienen de las señales o efectos que se han logrado reproducir en simulaciones basadas en ecuaciones físicas. El investigador intenta entender y reproducir un

fenómeno observado a través de las ecuaciones de la dinámica de fluidos y algunas ecuaciones físicas complementarias que representan simplificadamente los procesos reales por ejemplo la convección y la turbulencia en la capa límite planetaria. En la figura 3, se muestran algunos fenómenos identificados en las observaciones. Una vez que la descripción y simulación del fenómeno es suficientemente realista, se puede considerar una fuente de predictibilidad. Desde el punto de vista estadístico, el entender un fenómeno permite al pronosticador reducir el número de posibles estados futuros. Por ejemplo, si tenemos un evento El Niño apenas iniciando, el pronosticador pudiera anticipar aproximadamente las consecuencias conforme el evento se desarrolla y se disipa. La figura 4 es un compuesto del impacto en la temperatura y precipitación ante el fenómeno del Niño en diferentes regiones del globo. Es importante tener en cuenta que las conexiones estadísticas entre un fenómeno como el Niño con efectos en algunas regiones geográficas específicas es solo estadística. Es posible que los efectos de un evento El Niño se vean modificados por otros procesos ocurriendo a otras escalas de tiempo en otros dominios (Barnston et al., 2005) resultando en forzamientos locales o remotos. Además de El Niño existen docenas de procesos y mecanismos que influyen la variabilidad climática en una región específica. Los controles de gran escala son representados por índices o modos preferidos de variabilidad (Ver tabla en <http://rainbow.ideo.columbia.edu/~alexeyk/Papers/Kaplan2011picv.pdf>).



Prácticas actuales

Reconociendo la incertidumbre natural de los pronósticos, las prácticas actuales para realizar los pronósticos incluyen correr estos modelos en modo ensemble y organizando ensembles generados en varias instituciones en los llamados ensembles Multi-Modelos (Multi-Model Ensembles, MME; Kirtman et al., 2014). La idea de combinar las salidas de los modelos no es reciente; por décadas los pronosticadores han utilizado métodos ingeniosos para mezclar las salidas de un modelo con la de otro, o han seleccionado un modelo de un conjunto de modelos que tienen a su disposición. Se está dejando atrás la idea determinista de los pronósticos y una gran tarea es la de crear ensembles que representen la

Desarrollo y uso de sistemas numéricos de pronóstico estacional

Peña, M.

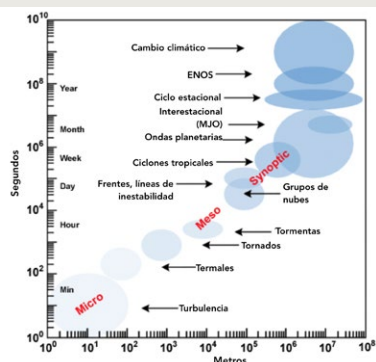


Figura 3. Escalas de tiempo y espacio de fenómenos climáticos. Imagen cortesía de The COMET Program, UCAR.

incertidumbre en el pronóstico. Lo que sí es nuevo es que ahora los pronosticadores se ven virtualmente inundados de cientos de salidas de modelos y no hay tiempo de analizarlos todos y mucho menos de combinarlos mentalmente en el corto tiempo que tienen para crear sus productos. Este problema se ha abordado a través de métodos de consolidación (e.g., Peña and van den Dool, 2008). Otros usuarios desearían poder utilizar estos ensembles para detectar solamente los casos extremos para tomar decisiones ante riesgos cuando valores de índices establecidos excedan umbrales específicos. Varios centros desarrollan algoritmos para el soporte de decisión (DSS), los cuales intentan sintetizar toda la información en índices críticos (e.g., Dutra et al., 2013) definidos por el usuario. La práctica de incorporar reanálisis y hindcasts como herramientas para post-procesar los datos crudos del pronóstico están logrando producir pronósticos probabilísticos estadísticamente confiables. Se ha notado que cuando los pronósticos promedio de los MME tienen mejor habilidad predictiva que los promedios de ensembles de modelos individuales y estos son mejores a los pronósticos deterministas del mismo modelo. Los MME también proveen un mayor muestreo de la distribución de probabilidad del pronóstico por lo que son útiles para la detección de extremos.

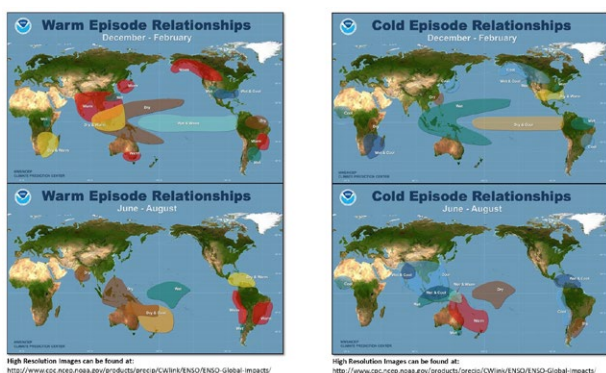


Figura 4. Compuesto del impacto global del ENOS en anomalías de temperatura y precipitación durante episodios calientes (El Niño) en los paneles de la izquierda y durante episodios fríos (La Niña) en los paneles de la derecha. Fuente: CPC/NWS/NOAA.

Conclusión y direcciones futuras

A medida que los modelos se vuelven más hábiles y confiables, los métodos para generar las guías numéricas del clima han pasado de enfoques subjetivos a técnicas

numéricas cada vez más objetivas y con menos supervisión.

Sin embargo, la intervención humana típicamente aumenta la habilidad y el valor de los productos finales y es esencial para la interpretación, optimización y uso efectivo del producto.

Los centros meteorológicos se esfuerzan por proporcionar productos informativos, precisos y oportunos incorporando e implementando nuevas ciencias y técnicas, lo que explica su estrecha relación con la comunidad científica. Los principales hitos científicos y tecnológicos asociados con la predicción del clima se han logrado a través de la colaboración nacional e internacional para atraer a expertos y fondos. Los desafíos para modelar de manera realista el sistema climático son muchos, pero aquellos destacados por la comunidad científica incluyen una mejor representación de modelos de procesos a escala fina en nubes, remolinos oceánicos e interacciones y retroalimentación superficiales, y una mejor integración de componentes climáticos modelo. Los pronósticos más hábiles también están condicionados a mayores recursos informáticos, sistemas de observación más extensivos, estrategias, asimilación efectiva de datos y esquemas de inicialización de modelos para los sistemas de predicción climática acoplados.

La gran cantidad de datos ahora disponible está dando curso a nuevos desarrollos interdisciplinarios como la introducción de conceptos tales como "Big Data", y "data-driven" models, lo cual plantea nuevas preguntas: ¿Cómo se incorporarán estas disciplinas dentro del esquema de un sistema numérico de pronóstico? O tal vez, ¿Cómo los sistemas numéricos de pronóstico se incorporarán en otros esquemas mayores por ejemplo sistemas social-ecológicos o población-naturaleza?

Referencias

- Bauer et al., 2015: The quiet revolution of numerical weather prediction. *Nature*, 525, 47-55.
- Barnston et al., 2005: Improving seasonal prediction practices through attribution of climate variability. *Bull. Atmos. Meteorol. Soc.*, 86, 59-72.
- Dutra et al., 2013: The extreme forecast index applied to seasonal forecasts. *ECMWF Technical Memorandum 703*.
- Gandin, L. S., 1963: Objective Analysis of Meteorological Fields. *Gidrometeorologicheskoe Izdaristvo (GIMZ)*, Leningrad. (Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.)
- Hoskins, B., 2013: Review Article.
- The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our science. *Q. J. R. Meteorol. Soc.* 139, 573-584.
- Kalnay E., 2002: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, December 2002
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 Seasonal-to-interannual prediction: Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, 95, 585-601
- Murphy, A. 1993: What is a good forecast?, *Wea. Forecasting*, 8, 281-292.
- Peña, M., and H. van den Dool, 2008: Consolidation of multimodel forecasts by ridge regression: Application to Pacific sea surface temperature. *J. Climate*, 21, 6521-6536
- Purser, J., 2011: Mathematical principles of the construction and characterization of a parameterized family of Gaussian mixture distributions suitable to serve as models for the probability distributions of measurement errors in nonlinear quality control. *NCEP Office notes* 468.
- Saha et al., 2014: The NCEP Climate Forecast System Version 2. *J. Climate*, 27, 2185-2208.
- Van den Dool 1994: Searching for analogues, how long must we wait?, *Tellus*, 46,314-324.