

Con técnica de big data investigadores predicen la evolución de la pandemia COVID-19

UNNE-CONICET

Un equipo interdisciplinario de la Universidad Nacional del Nordeste y el Conicet compuesto por físicos, matemáticos y bioquímicos, cuenta con resultados preliminares del procesamiento de datos de casos de COVID-19 y su predicción probabilística.

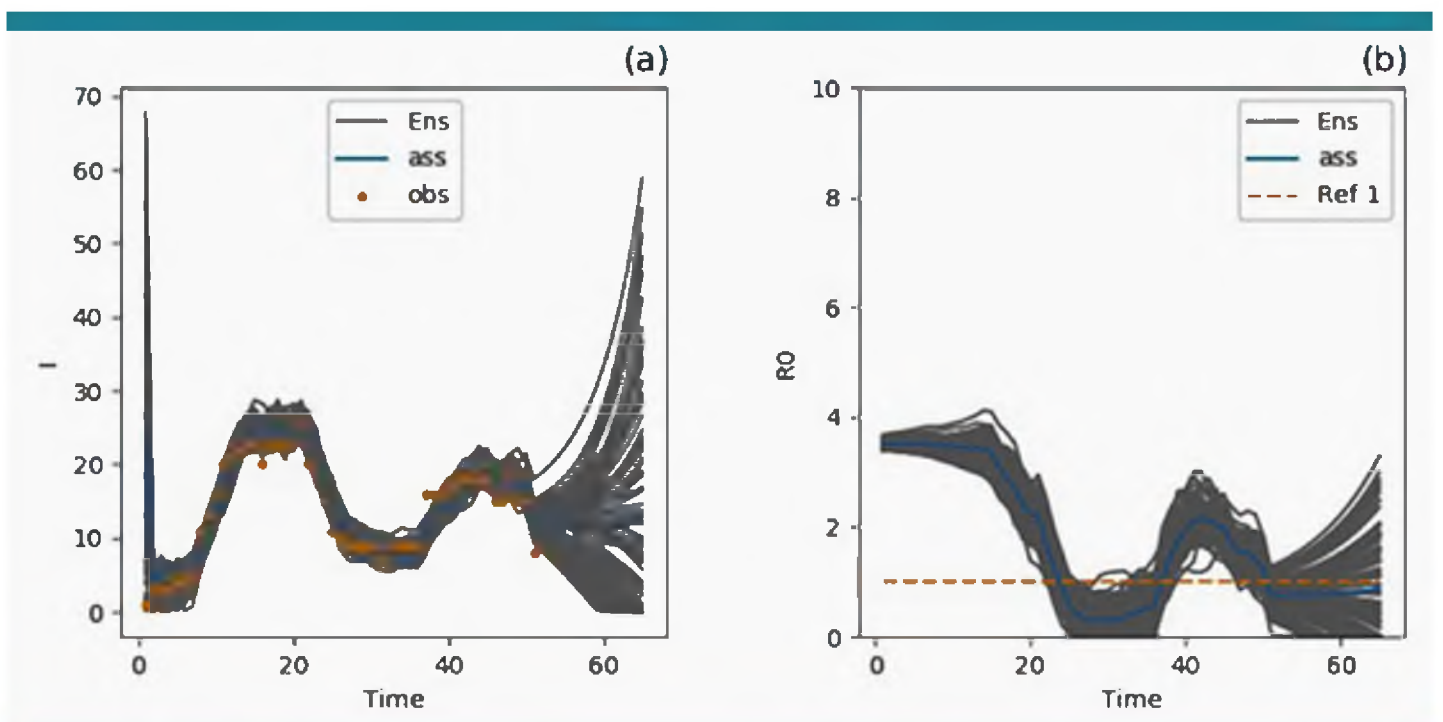
Este estudio aportará la proyección que tendrá la pandemia en cada una de las provincias. Utilizando una técnica de procesamiento de grandes datos (big data), denominada "asimilación de datos", se analiza la información que a diario es reportada de manera oficial (tanto los casos positivos como los decesos).

Una de las informaciones aportada por el trabajo, es el análisis del R cero ó tasa de contagio del coronavirus a nivel nacional. Antes del inicio de la cuarentena, la tasa era 3,5 es

decir un infectado contagiaba a 3,5 personas. La implementación de la cuarentena logró reducir considerablemente esa tasa 0,5, pero la flexibilización dispuesta en las últimas dos etapas volvió a incrementar R a valores entre 1 y 2.

La importancia de tener la estimación diaria de esta tasa es clave, ya que cuando se alcance un valor por debajo de 1, la epidemia se controla y el número de casos disminuirá con el tiempo. Por el contrario, si la tasa de contagio está por encima de 1, significa que la cantidad de infectados continuará creciendo.

Además de la tasa de contagio, la técnica de asimilación de datos brinda información estadística de las características de propagación del virus; los tiempos de incubación y de



La figura muestra la evolución de los casos positivos a partir del 20 de marzo en la provincia de Corrientes junto a las estimaciones que está realizando el sistema de asimilación. También se observan la predicción probabilística a dos semanas. El panel de la derecha muestra la tasa de contagios (R).

recuperación; la cantidad de enfermos severos que derivarán en hospitalizaciones o decesos. En un estudio reciente demostró, que la técnica también permite estimar la circulación del virus en personas asintomáticas, es decir el número de casos no-detectados por los testeos.

El grupo de trabajo que participa de este proyecto está liderado el doctor Manuel Pulido e integrado por los investigadores Tadeo Cocucci, Noelia Villafañe y Magdalena Lucini, de la Facultad de Ciencias Exactas y Naturales y Agrimensura y el CONICET. También participan referentes del Laboratorio Central de Corrientes, los bioquímicos Gerardo Andino, Natalia Ruiz Días y Viviana Gutnisky, responsables de los tests epidemiológicos en la provincia.

Paralelamente se conformó una red de colaboración internacional con distintos grupos que trabajan en asimilación de datos para COVID-19, particularmente con un grupo de investigadores de Francia coordinado por Pierre Tandeo del IMT Atlantique y de Inglaterra, con Alberto Carrassi director del Data Assimilation Research Centre, University of Reading, Inglaterra.

“La técnica de asimilación de datos también permite estimar la circulación del virus en personas asintomáticas”.

Ventajas. “Una de las grandes ventajas que posee la humanidad para poder controlar la actual pandemia COVID-19 comparada a las ocurridas en el pasado, es la gran disponibilidad de datos en tiempo real que se posee sobre la transmisión del virus en la población”, expresó el doctor Pulido, investigador del Departamento de Física de la FaCENA UNNE.

Esos datos al ser analizados en modelos matemáticos brindan información objetiva sobre por ejemplo, el impacto de las medidas de aislamiento social focalizadas o generales que permitan contener el virus. Los gobiernos disponen de esta manera de una herramienta en la toma de decisiones puntuales en base a un diagnóstico y con predicciones precisas sobre la evolución de la pandemia.

“Mientras mayor sea la información que se procese y más sofisticada la técnica de procesamiento de los datos mayor será el conocimiento de la situación de la pandemia”, señaló Pulido.

Asimilación de Datos. Dentro del procesamiento de grandes datos, las técnicas de asimilación se caracterizan por combinar en tiempo real información de dos fuentes: los

modelos dinámicos epidemiológicos y observaciones incompletas de distintas fuentes.

En un contexto donde la información no es completa ni tampoco exacta como es el caso del escenario planteado por el COVID-19, estas técnicas de asimilación permiten combinar toda la información circundante.

“La implementación de la cuarentena logró reducir considerablemente el Rcero a 0,5”

En el marco de la asimilación se pueden realizar predicciones o pronósticos probabilísticos, los cuales permiten obtener la incerteza de las predicciones y evaluar la probabilidad de eventos extremos. Los ejemplos donde se aplica la asimilación de datos son numerosos entre los que cabe citar los pronósticos meteorológicos, la robótica, vehículos autónomos, visión computarizada, seguimiento por GPS o radares, entre otros.

Los resultados de la aplicación de técnicas de asimilación en el brote de Wuhan China. en un trabajo liderado por Ruiyun Li del MRC Centre for Global Infectious Disease Analysis (Imperial College, Inglaterra), demuestran que estas técnicas tienen un enorme potencial para estimar la cantidad de casos no-detectados en la población basadas en observación parcial (casos de infectados detectados, movilidad, etc).

Antecedentes. El grupo de investigación de la UNNE-CONICET cuenta con antecedentes en el trabajo de técnicas para la asimilación de datos de modelos epidemiológicos.

El doctor Pulido recientemente desarrolló junto a Peter Jan van Leeuwen (Colorado State University) una técnica de asimilación de Monte Carlo secuencial que tiene muy buenos resultados en modelos epidemiológicos. Fue utilizado para un modelo de la dinámica del cólera en un trabajo que se publicó recientemente en el Journal of Computational Physics.

También en el grupo se está trabajando en la asimilación de datos de la dinámica de propagación del dengue. En este tema, Tadeo Cocucci, becario de CONICET, está realizando su tesis doctoral. Dados los devastadores efectos de la pandemia, el grupo se encuentra actualmente totalmente enfocado en la asimilación de datos de COVID-19.

Por Juan Monzón Gramajo