

JORNADAS DE SEMINARIOS 2021
-
DOCTORADO EN CIENCIAS EXACTAS E INGENIERIA FACET-UNT

Alumno: Jorge Habib Namour.

Graduado en: Licenciatura en Informática (FACET - UNT).

Directora: Dra. María Graciela Molina (FACET-UNT / CONICET).

Título de la Tesis: Modelo de predicción para meteorología del espacio usando técnicas de programación científica.

Lugar: Laboratorio de Computación Científica (LabCC), Departamento Ciencias de la Computación (DCC), Facultad de Ciencias Exactas y Tecnología (FACET), Universidad Nacional de Tucumán (UNT).

Año de inicio: Marzo 2019.



Actividad curricular → cursos

CURSO	C. HORARIA	ESTADO
Principios Generales sobre Radares y Aplicaciones en Geofísica Espacial	40	Cursando
Campo Magnético de la Tierra y Actividad Geomagnética	40	Aprobado
Análisis Bayesiano de Datos	40	Aprobado
Big Data Systems	15	Aprobado
Introducción a Machine Learning	40	Aprobado
Inteligencia Computacional para Problemas de Optimización Complejos	40	Aprobado
Regresión Logística	20	Aprobado
La Investigación Científica: Fundamentos y Metodologías	40	Aprobado
Introducción a las Misiones Espaciales, Producción de Satélites	20	Aprobado
Minería de Datos	64	Aprobado
Métodos Modernos de Series de Tiempo y sus Aplicaciones	50	Aprobado
Workshop en Técnicas de Programación Científica (WTPC)	70	Aprobado
Redes neuronales artificiales	60	Aprobado

Total hs. de cursos de postgrado aprobadas: 499/500

Total hs. de cursos específicos: 344

Total hs. de cursos generales: 155

Total hs. de cursos FACET-UNT: 290

Cumpliendo el requisito establecido en la resolución N°: 115/12: "...con un total de 500 horas obligatorias, el 70% de las cuales debe corresponder a asignaturas específicas y, por lo menos, 120 horas deben cumplirse por cursos dictados en la facultad."

Integrante en proyectos financiados

- “Aplicaciones de Meteorología del Espacio basadas en Inteligencia Artificial”. PIUNT E689 **2020-2022**.
Directora: Dra. María Graciela Molina.
- Estrategias de Computación Científica Aplicadas al Monitoreo de Eventos de Meteorología del Espacio. PIUNT **2018-2020**.
Directora: Dra. María Graciela Molina.
- Investigación de la Alta Atmósfera Terrestre Aplicada a Telecomunicaciones, Meteorología del Espacio y Cambio Climático. PICT-I-A-**2018**.
Director: Dr. Gustavo Mansilla.

Becas recibidas

- Otorgada por: CELFI (Centro Latinoamericano de Formación Interdisciplinaria) para asistir a la “COLAGE XI - Eleventh Latin American Conference on Space Geophysics” y a la “Latin American School on Data Analysis and Simulations for Space Physics”. Buenos Aires. **16/04/2018 al 26/04/2018**.

Artículos publicados

- “Consequences of a solar wind stream interaction region on the low latitude ionosphere: Event of 7 October 2015” (**2020**). Autores: M. G. Molina, S. Dasso, G. Mansilla, J. H. Namour, M. A. Cabrera, E. Zuccheretti. Lugar: Solar Physics. DOI: doi.org/10.1007/s11207-020-01728-7

Presentaciones en congresos y otras reuniones científicas

- “Consequences of a solar wind stream interaction region on the low latitude ionosphere: Event of 7 October 2015” **(2020)**. M. G. Molina, S. Dasso, G. Mansilla, J. H. Namour, M. A. Cabrera, E. Zuccheretti. European Space Weather Symposium (ESWS) 2020. Póster.
- “Técnicas y herramientas para la gestión de datos de un sistema de monitoreo de Meteorología del Espacio” **(2019)**. J. H. Namour, T. Torres Peralta, M. Ruiz, G. Amaya, G. Delorme, M. G. Molina. “Encuentro Científico de Investigadores de la FACET - ECIFACET“. FACET-UNT. Póster.
- “Space Weather Monitoring: challenges in data management” **(2018)**. J. H. Namour; T. Torres Peralta; M. Ruiz; M. G. Molina; M. Fagre ; F. Miranda Bonomi. “COLAGE XI - Eleventh Latin American Conference on Space Geophysics” y “Latin American School on Data Analysis and Simulations for Space Physics”. Centro Cultural de la Ciencia, Palermo, Buenos Aires y Universidad Nacional de Buenos Aires (UBA). Póster.
- “Addressing Space Weather Monitoring data handling problem” **(2017)**. M. G. Molina, J. H. Namour, T. Torres Peralta, M. Ruiz, M. A. Cabrera. Pydata 2017. San Luis. Argentina. Charla.

Trabajos aprobados para ser presentados:

- “Ionosphere F2 critical frequency forecasting using deep learning” **(2021)**. J. H. Namour, M. G. Molina, J. Ise. Applications of Statistical Methods and Machine Learning in the Space Sciences. Póster.
- “New Space Weather Services in Argentina: The Tucumán Space Weather Center”. 28 August – 4 September, **2021** – Sapienza Faculty of Engineering, Rome, Italy XXXIV General Assembly and Scientific Symposium (GASS) of the International Union of Radio Science (Union Radio Scientifique Internationale-URSI). Póster.

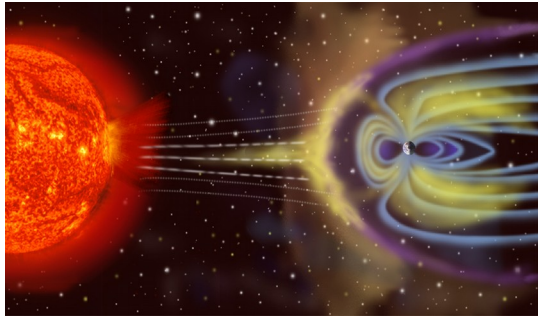
Asistencia a escuelas científicas

- “Web School of the Space Science on Dynamical Systems and Machine Learning Approaches to Sun-Earth Relations”. International School of Space Science (ISSS) L’Aquila, Italia. (virtual). **02/2021**.
- “Workshop: Machine Intelligence Training and applications to regional challenges”. Tucumán, Argentina. **03/2020**.
- “COLAGE XI - Eleventh Latin American Conference on Space Geophysics” y a la “Latin American School on Data Analysis and Simulations for Space Physics”. Universidad de Buenos Aires. **04/2018**.
- “Escuela de Ciencias Informáticas (ECI)”. Universidad de Buenos Aires. **07/2015 y 07/2017**.

Objetivo del trabajo

Estudiar y predecir la variabilidad de la ionosfera (f_0F_2 , TEC)

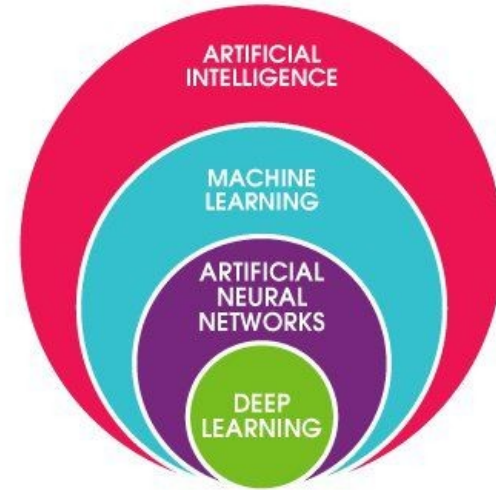
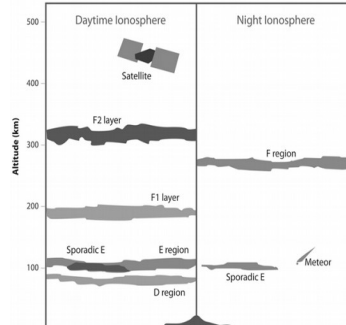
Meteorología del espacio

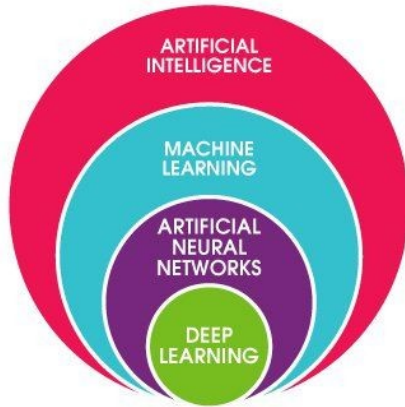


Modelo de predicción
basado en Machine Learning



Alta atmósfera terrestre





Aplicación: predicción de series de tiempo.

Tipo de aprendizaje: supervisado.

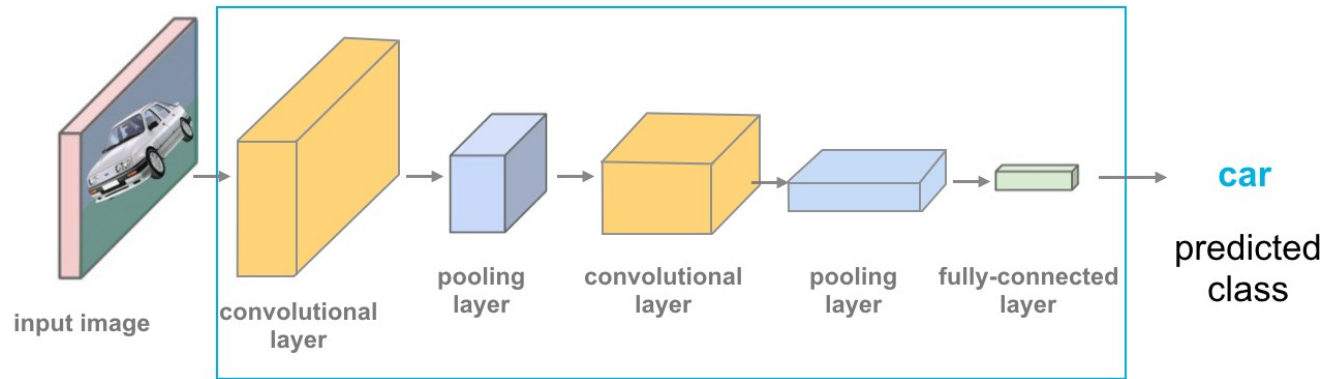
Pipeline:

- Recolectar datos.
- Preparar datos.
- Elegir el modelo. Configurar parámetros.
- Entrenar modelo.
- Evaluar modelo.
- Interpretar modelo.
- Generar reportes.

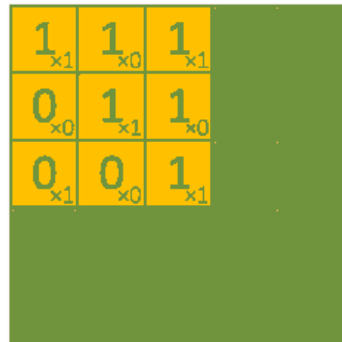
Algoritmos: Redes Neuronales Artificiales. Aprendizaje profundo.

- Redes Neuronales Convolucionales.
- Redes Neuronales Recurrentes (LSTM, GRU).

Redes Neuronales Convolucionales



Capa de convolución

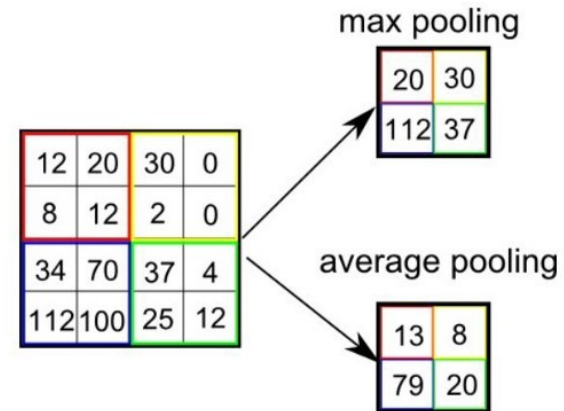


Image



Convolved
Feature

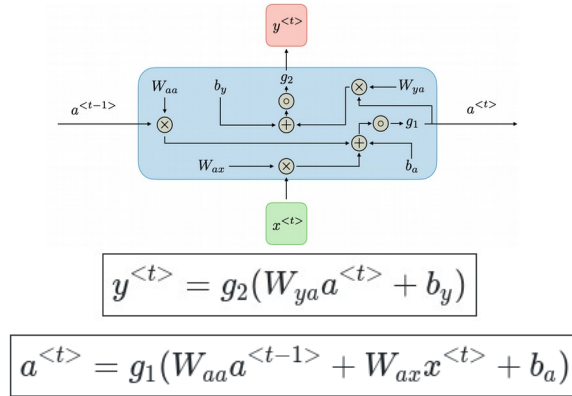
Capa de agrupación



Redes Neuronales Recurrentes

LSTM (Long Short Term Memory)/GRU (Gated Recurrent Unit)

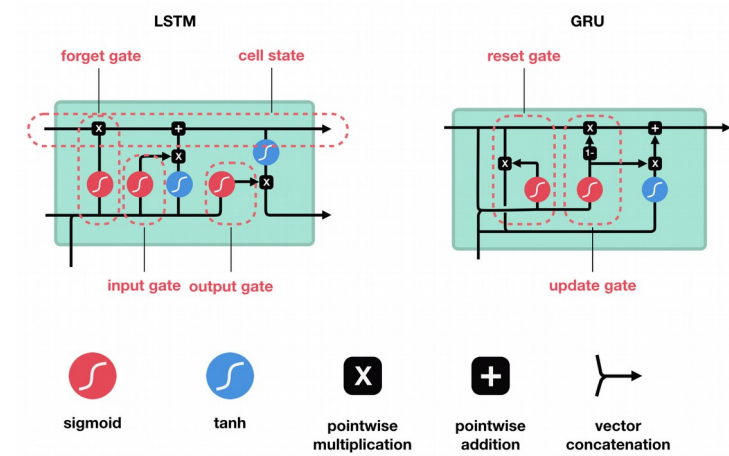
Red Neuronal Recurrente pura



Problema

Vanishing gradient $\left\| \frac{\partial h_i}{\partial h_{i-1}} \right\|_2 < 1$

Exploding gradient $\left\| \frac{\partial h_i}{\partial h_{i-1}} \right\|_2 > 1$



$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

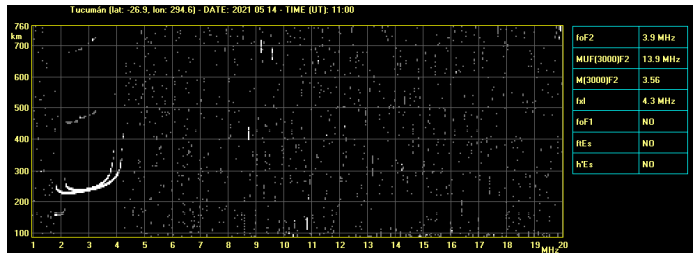
Análisis de series de tiempo univariable multipaso

Objetivo → Predecir a corto plazo de frecuencia crítica de la capa F2 (f0F2) sobre Tucumán.

[Fuente de datos]
Estación Ionosférica
Tucumán, Sondador
ionosférico Digital (AIS-
INGV)*

Tucumán (Lat: -26.9, Lon: 294.6) - DATE: 2021 05 14 - TIME (UT): 11:00
Autoscala 2007
foF2 03.9 MHz
MUF(3000)F2 13.9 MHz
M(3000)F2 03.56
fxI 04.3 MHz
foF1 NO
ftEs NO
h'Es NO

Modelo de archivo del Sondador ionosférico Digital (AIS-INGV)



Algoritmo empleado:
LSTM (Long Short-Term Memory).

Ejemplo de archivo final

```
yy,mm,dd,time,f0F2,f0F2_predicted,delta_f0F2,delta_porc_f0F2
2018,1,27,22:50,8.9,10.324946403503418,1.4249464035034176,0.1601063374722941
2018,1,27,23:00,8.6,10.2201089094543457,1.62010890945434574,0.1883940807686714
2018,1,27,23:10,8.7,10.118294715881348,1.4182947158813484,0.1630238113578718
2018,1,27,23:20,8.6,9.98796844482422,1.387968444824219,0.16139167963072315
2018,1,27,23:30,8.7,9.811396598815018,1.111396598815019,0.12774673549608262
2018,1,27,23:40,8.6,9.622149467468262,1.022149467468262,0.1188545892404956
2018,1,27,23:50,8.9,9.49406909942627,0.5940690994262602,0.06674933701418755
2018,1,28,00:00,8.8,9.32990837097168,0.529908370971679,0.06021686033769079
2018,1,28,00:10,8.8,9.1890230178833,0.3890230178833001,0.04420716112310228
2018,1,28,00:20,8.7,9.051101684570312,0.3511016845703132,0.0403565154678521
2018,1,28,00:30,8.9,8.905959129333496,0.0059591293334957385,0.0006695650936512067
2018,1,28,00:40,9.1,8.748976707458496,-0.35102329254150355,-0.03857398819137402
2018,1,28,00:50,9.1,8.663669586181639,-0.436330413818359,-0.0479483971228966
2018,1,28,01:00,9.3,9.380717277526855,0.08071727752685476,0.008679277153425242
2018,1,28,01:10,9.1,9.368764877319336,0.2687648773193363,0.029534601903223767
2018,1,28,01:20,8.8,9.320198059082033,0.5201980590820305,0.059113415804776186
2018,1,28,01:30,8.6,9.230523109436037,0.6305231094360355,0.07331664063209715
2018,1,28,01:40,8.1,9.1576500475709,1.0576500475708011,0.1305741910581236
2018,1,28,01:50,7.3,9.10860596460645,1.808605964606447,0.24658351323376
2018,1,28,02:00,6.6,8.986719131469727,2.3867191314697274,0.3616241108287465
2018,1,28,02:10,6.3,8.852682113647461,2.552682113647461,0.4051876370868986
2018,1,28,02:20,5.8,8.706656455993652,2.9066564559936525,0.5011476648264919
2018,1,28,02:30,5.3,8.535666384277344,2.235666384277344,0.6104917706183668
2018,1,28,02:40,5.6,8.336467742919924,2.7364677429199222,0.4886549540928433
2018,1,28,02:50,5.6,8.198328018188478,2.5983280181884774,0.4639871461050852
2018,1,28,03:00,5.8,8.052109718322754,2.252109718322754,0.3882947790211645
2018,1,28,03:10,5.4,7.9408607482910165,2.5408607482910157,0.4705297682020398
```

Características de los datos con los que se opera:

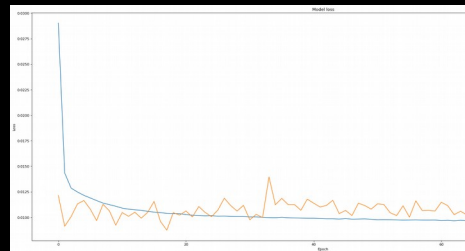
Rango total de datos: 01/09/2007 – 29/12/2017

Cantidad de datos del rango: 712.790

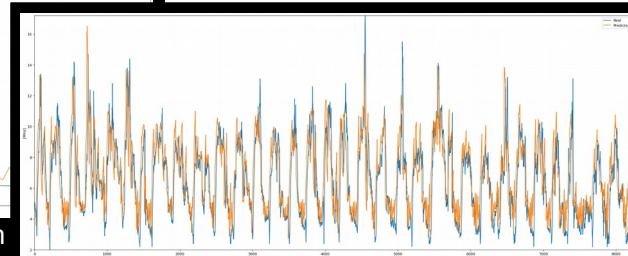
Resolución: 10 minutos (excepto en campañas -5 minutos).

Cantidad de datos faltantes (NaN): ~%45

Algunos resultados



Entrenamiento vs. Validación



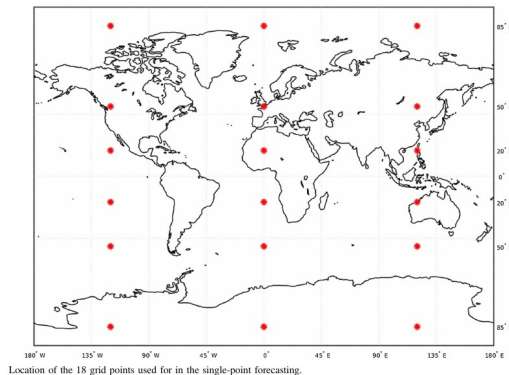
Valor real vs. Valor predicho

* Estación Ionosférica Tucumán. Sondador ionosférico Digital (AIS-INGV). Proyecto en colaboración con Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italia, Facultad Regional Tucumán, Universidad Tecnológica Nacional. Agosto 2007 a la fecha. Resp. Científicos: Dr. Enrico Zuccheretti (INGV), Dr. Miguel A. Cabrera, Dr. Rodolfo G. Ezquer, Dr. Luis A. Scida. Bajo responsabilidad de Laboratorio de Telecomunicaciones. FACET-UNT.

* Trabajo en conjunto con colegas del: Abdus Salam International Centre for Theoretical Physics (ICTP)

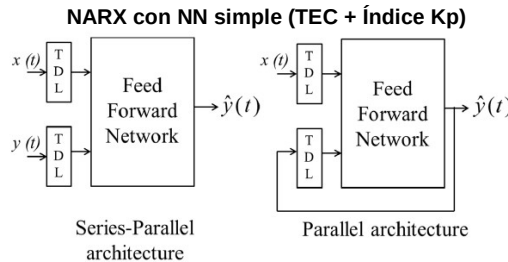
Análisis de series de tiempo multivariable multipaso

Objetivo → Predecir 24 hs. de TEC en diferentes puntos de una grilla (TEC + Índice Kp).



Location of the 18 grid points used for in the single-point forecasting.

Actualmente



Algoritmos a emplear:

- Redes Neuronales Recurrentes.
 - Long Short-Term Memory (LSTM).
 - Gated Recurrent Units (GRU).
- Redes Neuronales Convolucionales.

```
YYYYMMDD Kp[8] Sum ap[8] Ap
19970101 2 2-1-1-1 1 1 1 9+ 7 6 3 3 4 4 4 5 4
19970102 1+1 1+2-2-2-1 1+1 5 4 5 6 6 6 6 4 5
19970103 1+0+1 0+1 1+1 5 3 2 4 2 4 5 4 4
19970104 1+1 1 1+0+0+1 7- 5 4 4 5 2 2 2 4 4
19970105 1+1+1 1 1+1-0+ 7+ 5 5 3 4 3 5 3 2 4
19970106 1 1-0+1 1+1-0+0+ 5 3 3 2 3 5 3 2 2 3
19970107 2+4 2 2+3-3 2+2-20 9 27 7 9 12 12 9 6 11
19970108 3-3+1+1 2 1+1+2-15- 12 18 5 4 7 5 5 6 8
19970109 2+2+1 1-0+1+2 11 7 5 9 4 3 2 5 7 5
19970110 3-5-6 5 4-3+4 3-32 12 39 80 48 22 10 27 12 32
19970111 5-4 2+4 4-1+2-1+23 39 27 9 27 22 5 6 5 18
19970112 2-2 5-3-2 3-2+2 20 6 7 39 12 7 12 9 7 12
19970113 2 3-2-1 2+2-2-1 14 7 12 6 4 9 6 6 4 7
19970114 1 1-1+1-1+1-2-2- 9 3 5 3 5 4 6 6 4
19970115 1-1-1 1 2-2 1 0+ 8+ 4 3 3 4 6 7 4 2 4
19970116 1 1 1-1-0+0 0+1 5 4 4 3 3 2 0 2 4 3
19970117 0 0 0 0 1-1 1+ 4 0 0 0 0 3 4 4 5 2
19970118 3 2 1+1+1-1-0+10 15 7 5 5 9 3 3 2 5
19970119 1 1-0+1 2 3 1+1 10+ 4 3 2 4 7 5 6 6
19970120 1-0+1+1 2-3 2 3-13- 3 2 5 4 6 15 7 12 7
19970121 3+2-2-2 3 3 2 1 18- 18 6 6 7 15 15 7 4 10
19970122 2+1 1 1-2-2-2-12 7 9 4 4 3 6 6 6 6
19970123 2+0+1-1 1-1-0+ 7- 9 2 3 3 4 3 3 2 4
19970124 0+1 1-1-1+1 2+ 8 2 4 3 3 3 5 4 9 4
19970125 2+1+1 1 1 0+2+3-12 9 5 4 4 4 2 9 12 6
19970126 2+2 1-2-4-4-4+5-23 9 7 3 6 22 22 32 39 18
19970127 3-4-2 3-1+2+3+4 23+ 15 22 7 12 5 18 10 27 16
19970128 3+5-3-3-4 4 4+4-29+ 18 39 12 12 27 32 22 24
19970129 3-2-2-2 2-3 2+17+ 12 7 6 7 7 7 6 15 9 9
19970130 3+3-1+3 4 4+3+23+ 15 9 12 5 15 27 22 18 15
19970131 3-2 1-1 1-2 2+14+ 12 15 3 3 4 6 7 9 8
19970201 2 1-1-1 2+2 1-1-10 7 3 3 4 9 7 3 3 5
19970202 2-2+2+1 2-3+3-19- 6 9 9 4 6 18 18 15 11
19970203 4-1 1-1-0+1 1+3-11+ 22 4 3 3 2 4 5 12 7
19970204 1-0+1 2 1+1-1-1 8- 3 2 4 7 5 3 3 4 4
```

Datos del índice Kp

```
1.0 IONOSPHERE MAPS MIX IONEX VERSION / TYPE
tecms2ionex 2.swk gAGE/UPC PCN / RUN BY / DATE
Global Ionosphere maps for day 300, 2000 (26-10-2000) DESCRIPTION
IONEX file containing the COMBINED IGS Ionosphere maps from the combination of the following IAAcs: DESCRIPTION
codg DESCRIPTION
emrg DESCRIPTION
esop DESCRIPTION
jplg DESCRIPTION
upcg DESCRIPTION
Contact address: Manuel Hernandez-Pajares DESCRIPTION
Res. group of Astronomy and Geomatics DESCRIPTION
Technical Univ. of Catalonia (gAGE/UPC) DESCRIPTION
Mod. C3 Campus Nord UPC DESCRIPTION
E08034-Barcelona, SPAIN DESCRIPTION
e-mail: manuel@mat.upc.es DESCRIPTION
2000 10 26 1 0 0 EPOCH OF FIRST MAP
2000 10 26 23 0 0 EPOCH OF LAST MAP
7200 INTERVAL
12 # OF MAPS IN FILE
COS2 MAPPING FUNCTION
0.0 ELEVATION CUTOFF
Combined TEC calculated as weighted mean of input TEC valuesOBSERVABLES USED
0 # OF STATIONS
6371.0 # OF SATELLITES
2 BASE RADIUS
2 MAP DIMENSION
450.0 450.0 0.0 HGT1 / HGT2 / DHGT
87.5 -87.5 -2.5 LAT1 / LAT2 / DLAT
-180.0 180.0 5.0 LON1 / LON2 / DLON
EXPONENT
TEC values in 0.1 TECUs; 9999 if no value available COMMENT
DCB values in nanoseconds, reference is Sum of SatDCBs = 0 COMMENT
DIFFERENTIAL CODE BIASES START OF AUX DATA
01 -0.952 0.147 PRN / BIAS / RMS
02 -2.169 0.165 PRN / BIAS / RMS
03 -0.317 0.141 PRN / BIAS / RMS
04 0.771 0.090 PRN / BIAS / RMS
05 -0.420 0.108 PRN / BIAS / RMS
06 -0.175 0.115 PRN / BIAS / RMS
07 -1.987 0.134 PRN / BIAS / RMS
08 -0.366 0.081 PRN / BIAS / RMS
09 0.261 0.184 PRN / BIAS / RMS
10 -1.748 0.104 PRN / BIAS / RMS
11 2.358 0.096 PRN / BIAS / RMS
13 4.372 0.133 PRN / BIAS / RMS
15 -1.392 0.314 PRN / BIAS / RMS
17 -1.730 0.188 PRN / BIAS / RMS
19 -1.152 0.092 PRN / BIAS / RMS
20 1.072 0.149 PRN / BIAS / RMS
21 -1.671 0.130 PRN / BIAS / RMS
22 -0.447 0.164 PRN / BIAS / RMS
23 -1.285 0.142 PRN / BIAS / RMS
24 -2.487 0.095 PRN / BIAS / RMS
25 1.673 0.107 PRN / BIAS / RMS
26 0.708 0.173 PRN / BIAS / RMS
27 -0.338 0.083 PRN / BIAS / RMS
28 3.714 0.089 PRN / BIAS / RMS
29 1.361 0.099 PRN / BIAS / RMS
30 1.732 0.176 PRN / BIAS / RMS
31 0.612 0.116 PRN / BIAS / RMS
DIFFERENTIAL CODE BIASES END OF AUX DATA
END OF HEADER
1 START OF TEC MAP
2000 10 26 1 0 0 EPOCH OF CURRENT MAP
87.5-180.0 180.0 5.0 450.0 LAT1/LON1/LON2/DLON/H
210 207 190 173 171 168 166 164 162 165 167 161 156 154 167 213
231 235 236 239 242 242 239 230 222 218 213 206 199 194 192 189
186 175 166 164 164 166 166 166 164 161 159 150 160 157 155 155
157 156 158 163 164 163 162 159 144 144 151 160 169 179 189 195
166 189 180 173 173 184 196 205 210
```

Modelo de archivo IONEX. IONEX VERSION 1.0
gdc.cddis.eosdis.nasa.gov/

Próximos pasos

- Actualmente en el proceso de escritura de un segundo paper mostrando los resultados preliminares de lo expuesto anteriormente.
- Finalizar el requerimiento de cursos de postgrado.
- Asistencia a eventos internacionales.



<https://spaceweather.facet.unt.edu.ar/>

Tucumán Space Weather Center (TSWC) es un programa de la Universidad Nacional de Tucumán. El centro surge a partir del trabajo conjunto del:

- Laboratorio de Computación Científica (LabCC <https://catedras.facet.unt.edu.ar/labcc/>).
- Laboratorio de Telecomunicaciones (LabTel <https://catedras.facet.unt.edu.ar/labtel/es/catedra/>).
- Laboratorio de Ionósfera, Atmósfera Neutra y Magnetósfera (LIANM).

MUCHAS GRACIAS