

## Automated Detection of Solar Flare-Induced Ionospheric Disturbances via Hidden Markov Models and VLF Propagation

A. Arriola<sup>1</sup>, C. Castromonte<sup>2</sup>, L. Otiniano<sup>1</sup>

<sup>1</sup> Comisión Nacional de Investigación y Desarrollo Aeroespacial (CONIDA)

<sup>2</sup> Universidad Nacional de Ingeniería

### 1 Propagación VLF

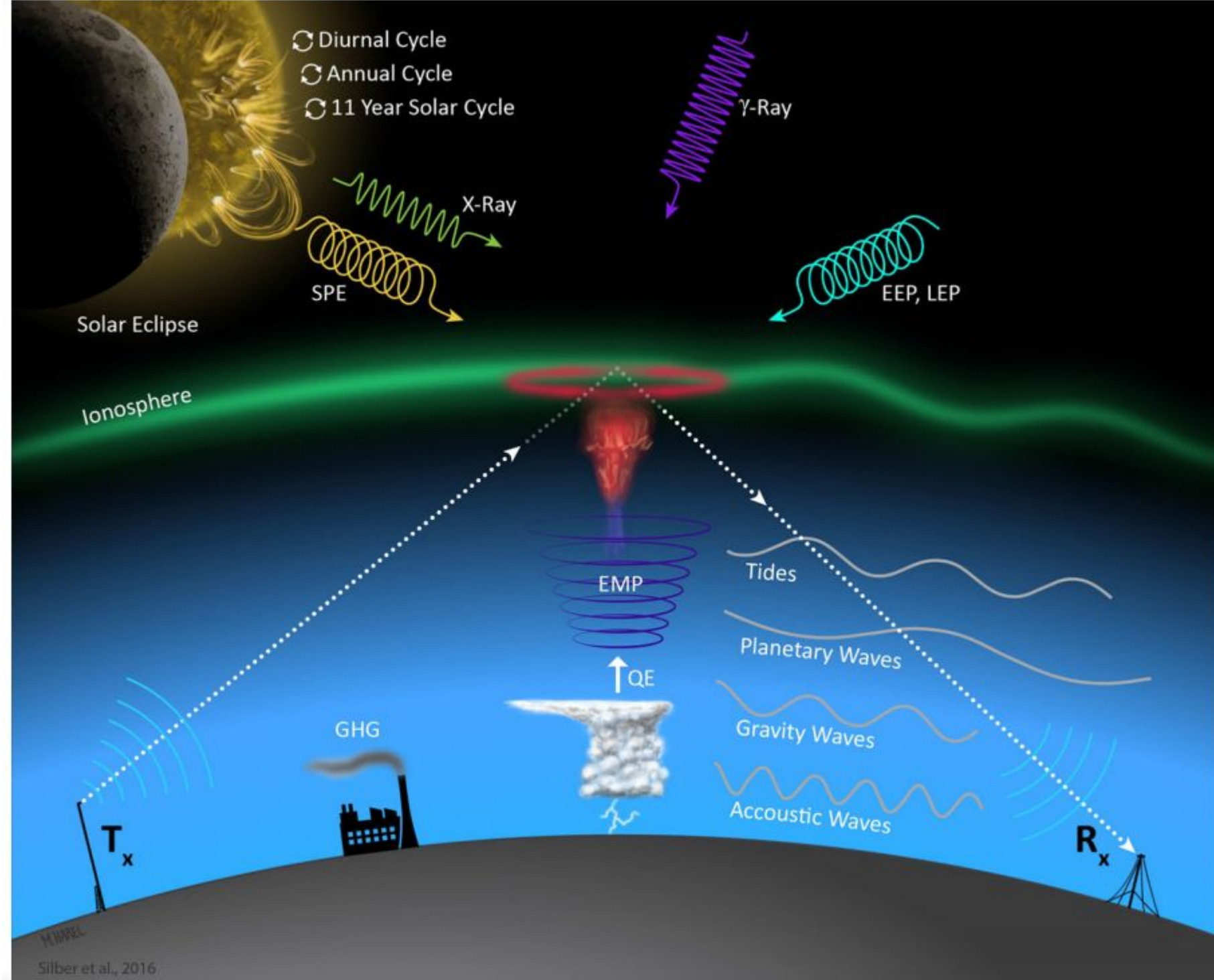


Figura 1. Propagación de ondas VLF y la interacción Sol-Ionosfera.

Las onda VLF se propagan largas distancias mediante a través de la guía de onda tierra ionosfera. Sin embargo esta propagación se ve afectada por alteraciones en la capa inferior de la ionosfera (region D). Estas alteraciones son producto de emisiones solares energéticas. Ver figura 1 (Silber, 2016).

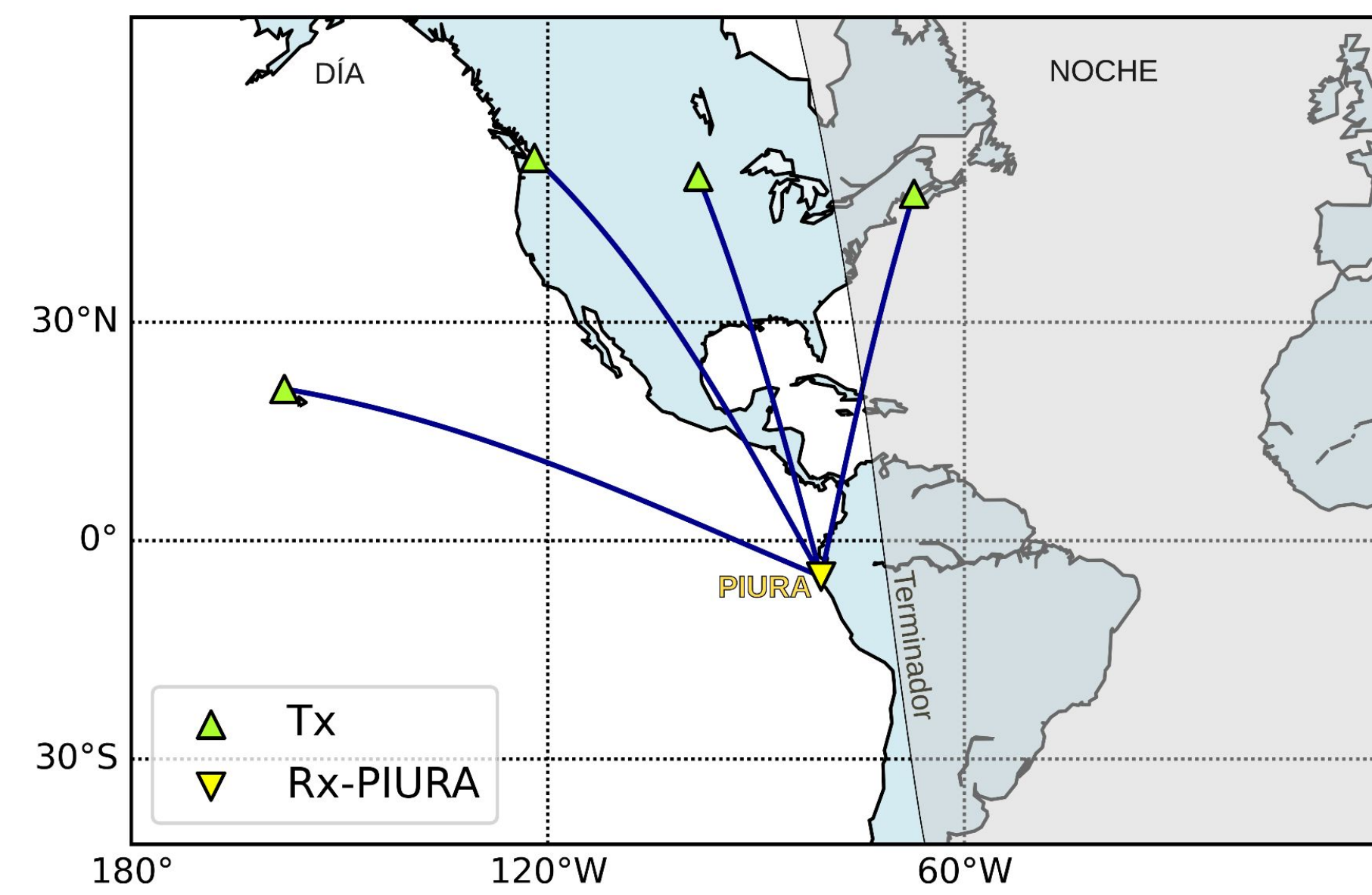


Figura 2. Great Circle Paths entre las Transmisoras (Tx) y la receptora VLF en Piura. Se observa el paso del terminador y su intersección con el trayecto de propagación.

Para la obtención de datos se empleó una receptora VLF instalada en Piura. Esta registra principalmente señales provenientes del hemisferio Norte. El paso del terminador es el principal modulador de las señales debido a la alteración en la concentración de carga en la ionósfera.

### 2. Hidden Markov Models (HMM)

Un Modelo Oculto de Markov se basa en la especulación de que existen estados subyacentes (no visibles) a un conjunto de observaciones. Estos estados experimentan un transición entre ellos en función de probabilidades de transición de los estados anteriores.

La probabilidad de que  $X_{n+1}$  esté en el estado  $j$  dado que  $X_n$  está en el estado  $i$  es la probabilidad de transición en un paso de  $i$  a  $j$  y la denotaremos como  $P_{ij}^{n,n+1}$ , tal que

$$P_{ij}^{n,n+1} = P(X_{n+1}=j | X_n=i)$$

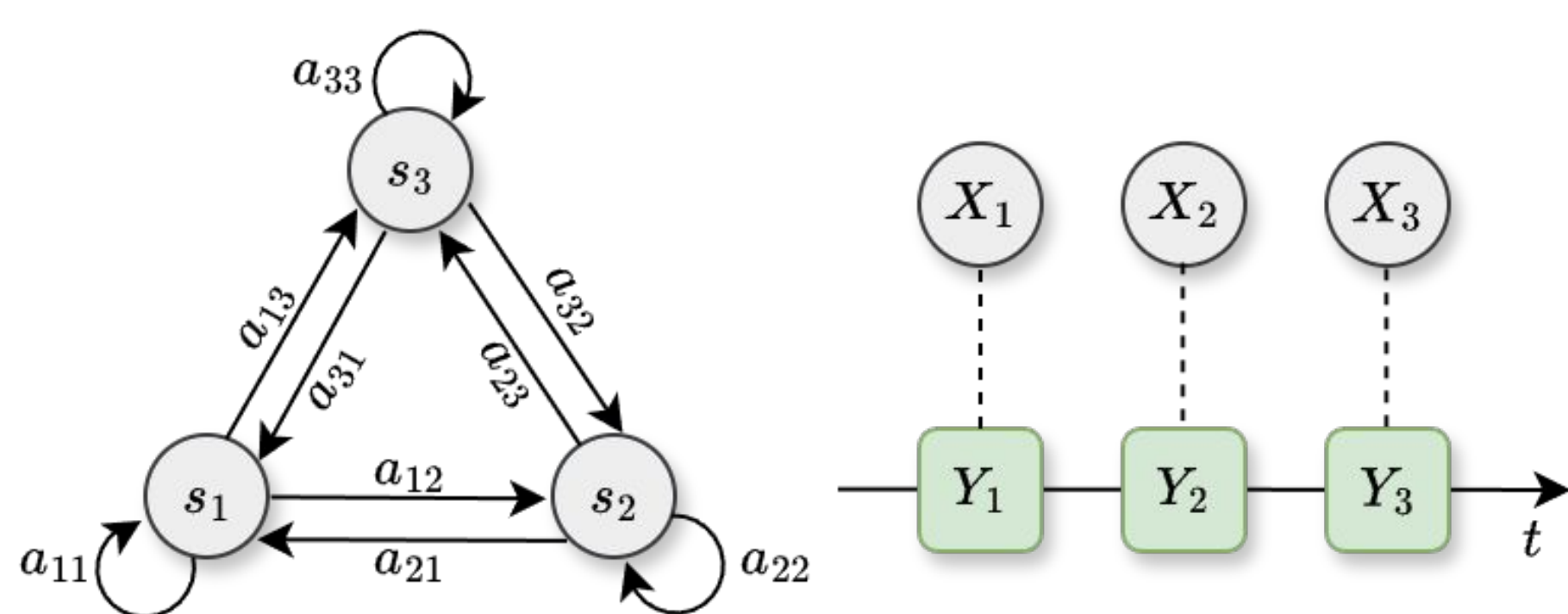


Figura 3. A la izquierda: Diagrama de transición de estados  $s_i$ . A la derecha: Observaciones  $Y_i$  y sus respectivos estados  $X_i$  como una secuencia temporal.

Adicionalmente, el modelo también contiene las probabilidades de emisión que modelan la probabilidad de observar  $Y_i=k$  dado que el sistema se encuentra en el estado  $i$ . Y también, el modelo cuenta con un distribución inicial de los estados ocultos (Rabiner, 1989).

Agradecimientos: Este trabajo ha sido financiado por CONCYTEC a través de su programa PROCENCIA en el marco del concurso “Proyectos de Investigación Básica 2023-01”, según contrato No. PE501082050-2023-PROCENCIA.

### 3. Metodologia

Para el modelado del sistema se etiquetaron las perturbaciones causadas por explosiones solares o *solar flares* en la señal VLF. Estas explosiones están correlacionadas con perturbaciones en el flujo de rayos X. Las perturbaciones se caracterizan con dos estados: *onset* (estado 1) y *decay* (estado 2). El comportamiento normal de la señal la consideramos con estado 0.

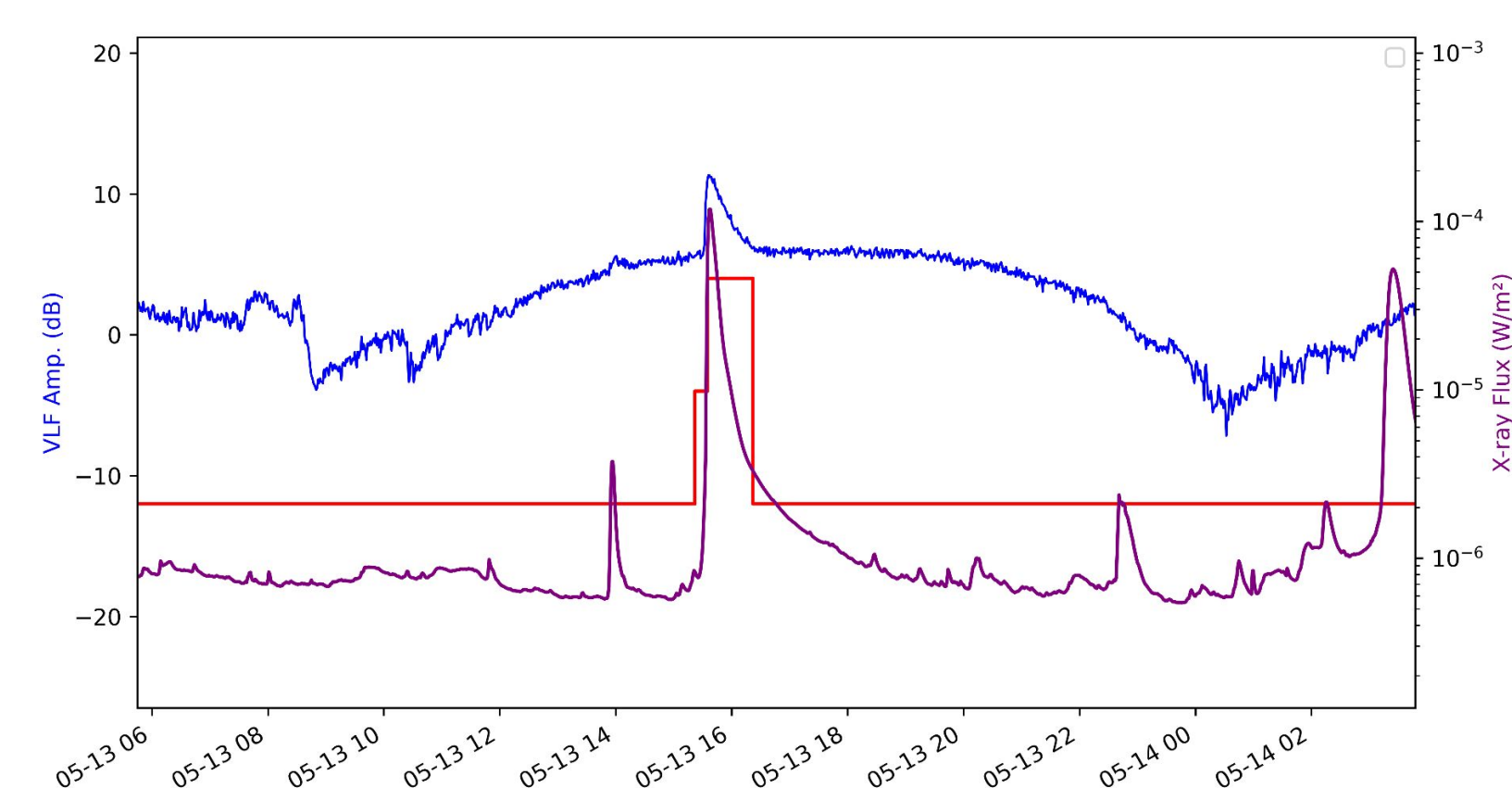


Figura 4. Caracterización de un solar flare como perturbación en la señal VLF. Se muestra la correlación entre el flujo de Rayos y la amplitud. Asi mismo, en rojo se indican los estados asignados.

Debido a la naturaleza de la señal fue necesario realizar una descomposición de la señal haciendo un preprocesamiento removiendo el background y haciendo uso de la derivada (rapidez) de la señal.

A partir de los estados etiquetados se calcula la matriz de probabilidades transición, y con este valor se ajusta el modelo usando el algoritmo EM (Expectation Maximization).

### 4. Resultados

Componentes de la señal

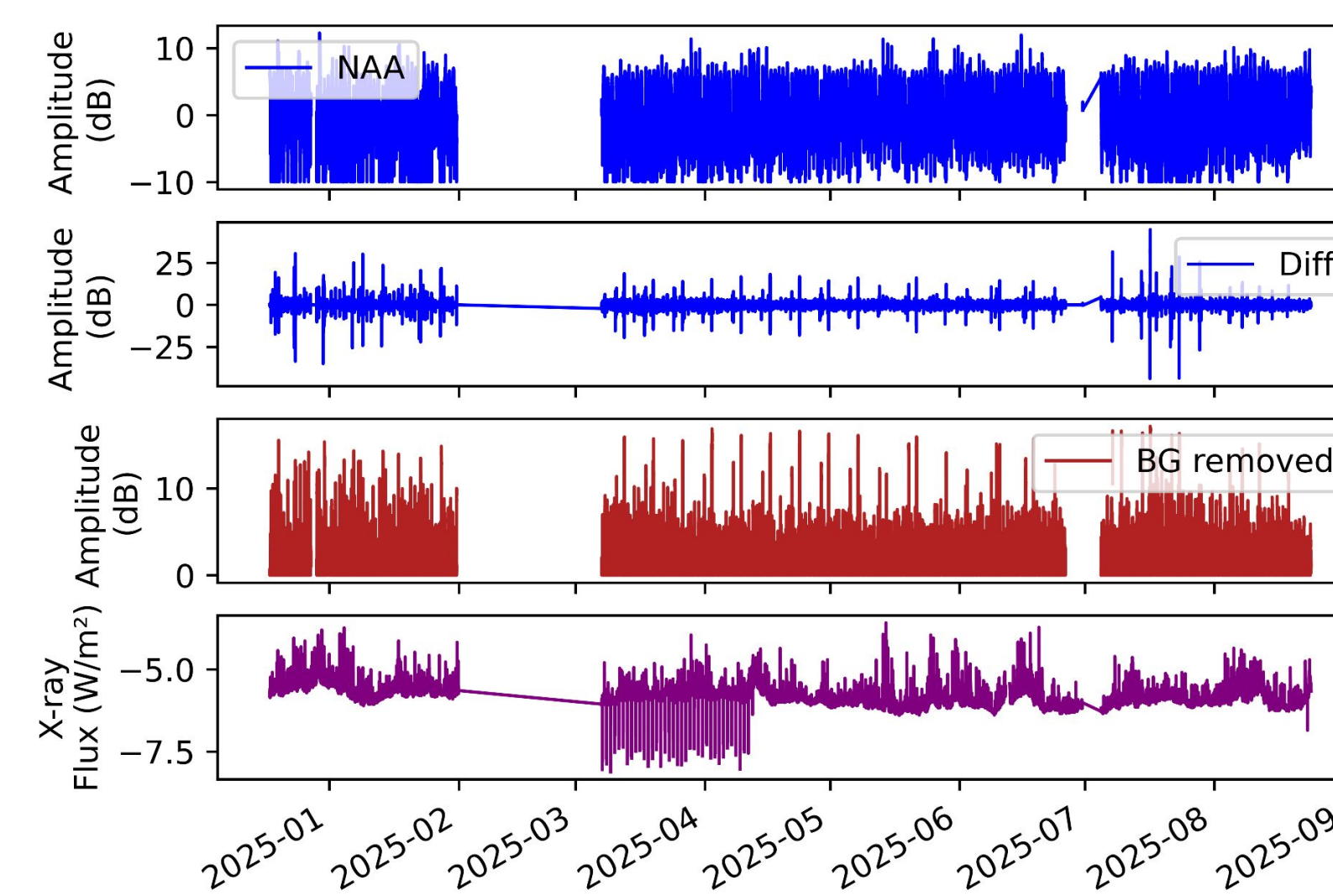


Figura 5. Componentes de la señal par como variables de ingreso para el ajuste del modelo.

Matriz de transición

Estados	Normal	Onset	Decay
Normal	0.999	0.574	0
Onset	0	0.936	0.063
Decay	0.031	0	0.968

**Rendimiento del modelo**

Con el modelamiento ya las condiciones establecidas en la metriz de transición, el modelo alcanzó una precisión del 76% con un recall del 74% para la detección del estado **Onset**. Este estado es de interés por ser el que nos puede brindar la detección anticipada.

Estados	Normal	Onset
Normal	0	8
Onset	9	25

### 5. Conclusions

El modelo logra clasificar el 76% de los casos de solar flares presentes en el ultimo 20% del dataset con un recall del 74%. Esto es un indicador de la viabilidad del modelo para este tipo de tareas. Sin embargo, existen otro tipo de implementaciones que podrían mejorar el rendimiento, uno de estos casos puede ser el uso de emisiones de Mezcla Gaussianas el cual permite modelar con mayor fidelidad la densidad de probabilidad de las emisiones, capturando mejor la complejidad de los datos que no se ajustan estrictamente a una única campana de Gauss.