

**IEEE**

Sección Uruguay



**URUCON2024**

35° ANNIVERSARY OF  
IEEE URUGUAY

**IEEE URUCON**

INNOVACIÓN Y  
TECNOLOGÍA PARA  
UN FUTURO MEJOR

Montevideo, Uruguay  
18 to 20, November



Aprendizaje por refuerzo aplicado al despacho óptimo de sistemas de energía eléctrica con alta integración de energías renovables variables y sistemas de almacenamiento

Ruben Chaer

# Plan de la presentación

- 1) **SimSEE**. Plataforma de Simulación de Sistemas de Energía Eléctrica especialmente diseñada para sistemas con alta integración de energías renovables.
- 2) **El Tractorcito**. Aprendizaje por refuerzo aplicado al despacho energético.
- 3) **OddFace+HC**. Optimizador de Inversiones con Herencia Cultural.

PARTE 1)

Plataforma de Simulación de Sistemas de Energía Eléctrica



Necesidades impuestas por las ERNC

- + Orientado a Objetos
- + Demanda Neta
- + Postizado Dinámico
- + Modelos CEGH



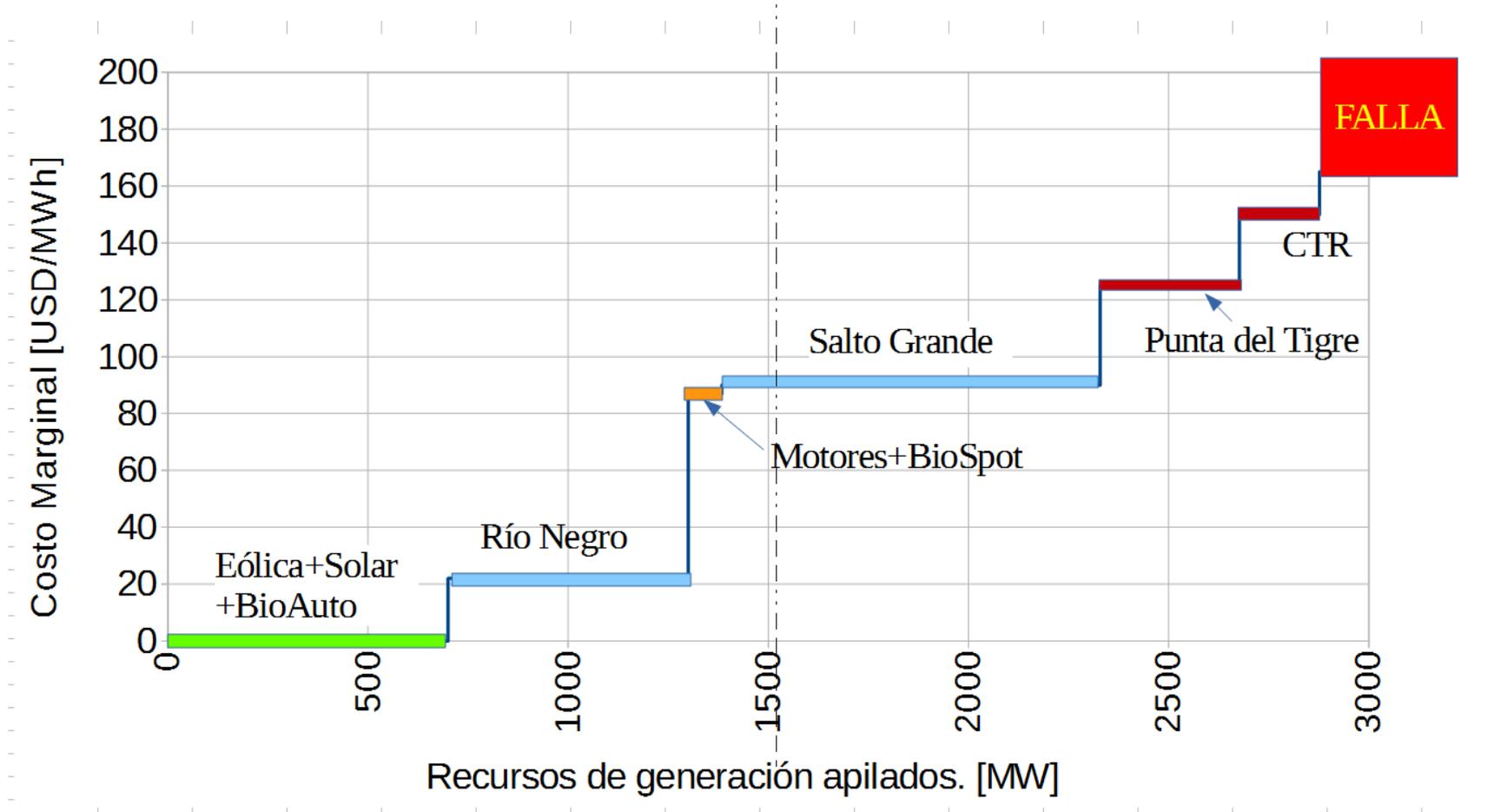
¿Por qué y para qué?



El objetivo último es abastecer La Demanda de energía al menor costo posible en condiciones de calidad y seguridad adecuadas.



# Orden de mérito y Costo Marginal. Solo Costos Variables

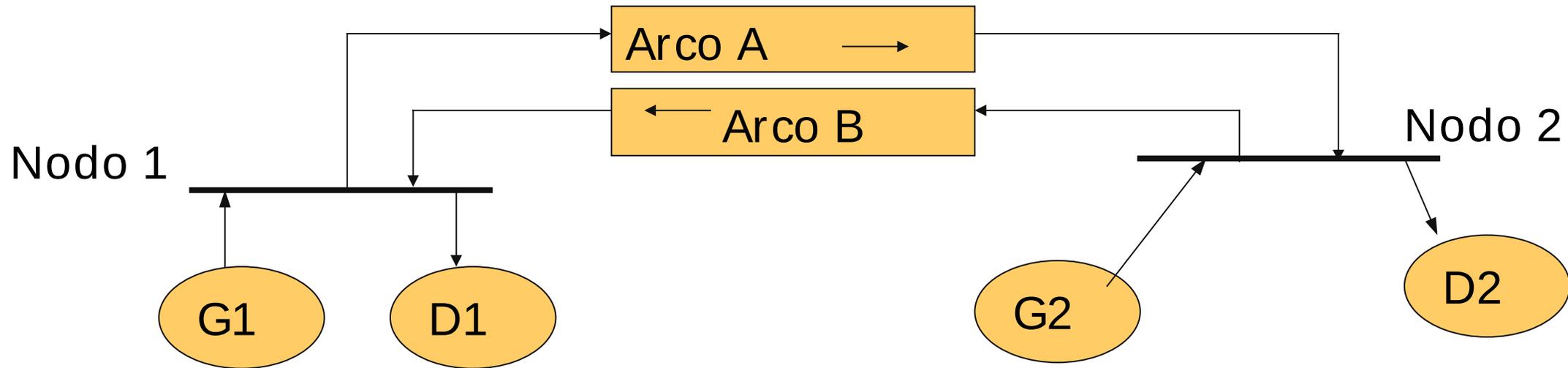


# Horizontes temporales de estudio y operación

- **mili-segundos:** Estatísticos de los generadores - control distribuido / Protecciones.
- **segundos:** Regulación de Frecuencia - consignas a generadores.
- **minutos:** Reserva Operativa Rodante (2.1% de la Demanda)
- **10 a 20 minutos:** Reserva Flexible (2.1% de la Demanda)
- **>= 20 minutos:** Reserva Fría. (3% de la Demanda; 20 minutos)
- **horas:** Seguimiento de la Demanda Horaria
- **días:** Programación del uso de los recursos
- **meses:** Optimización embalses y gestión de combustibles
- **años:** Optimización de inversiones

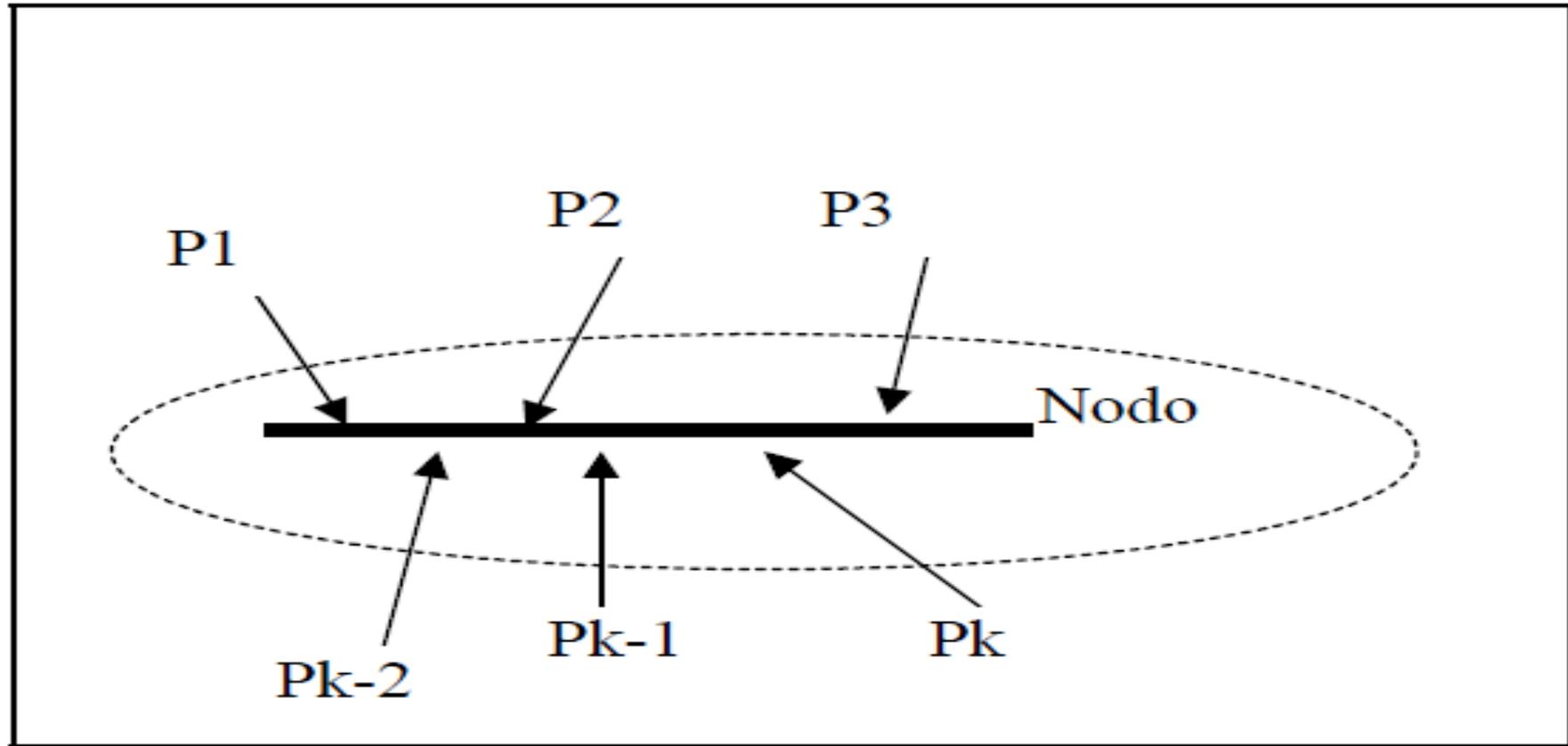


# Nodos, Arcos, Generadores y Demandas

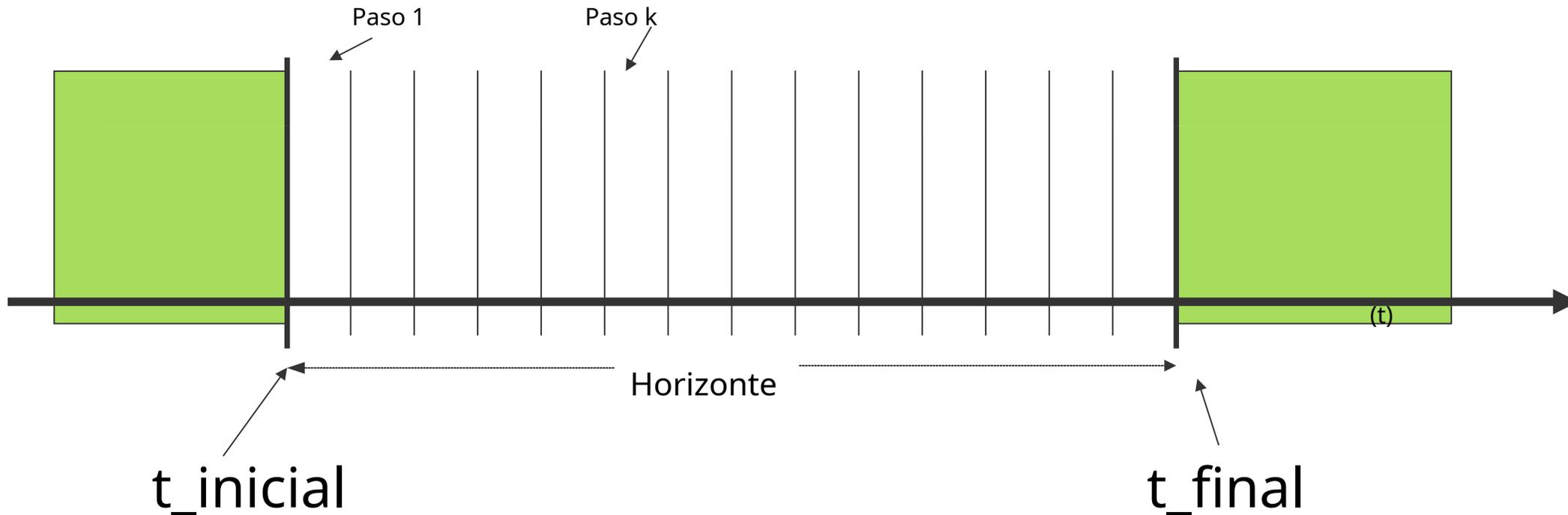


# Despacho y restricciones de Nodo

$$\sum_k P_k = 0 ;$$



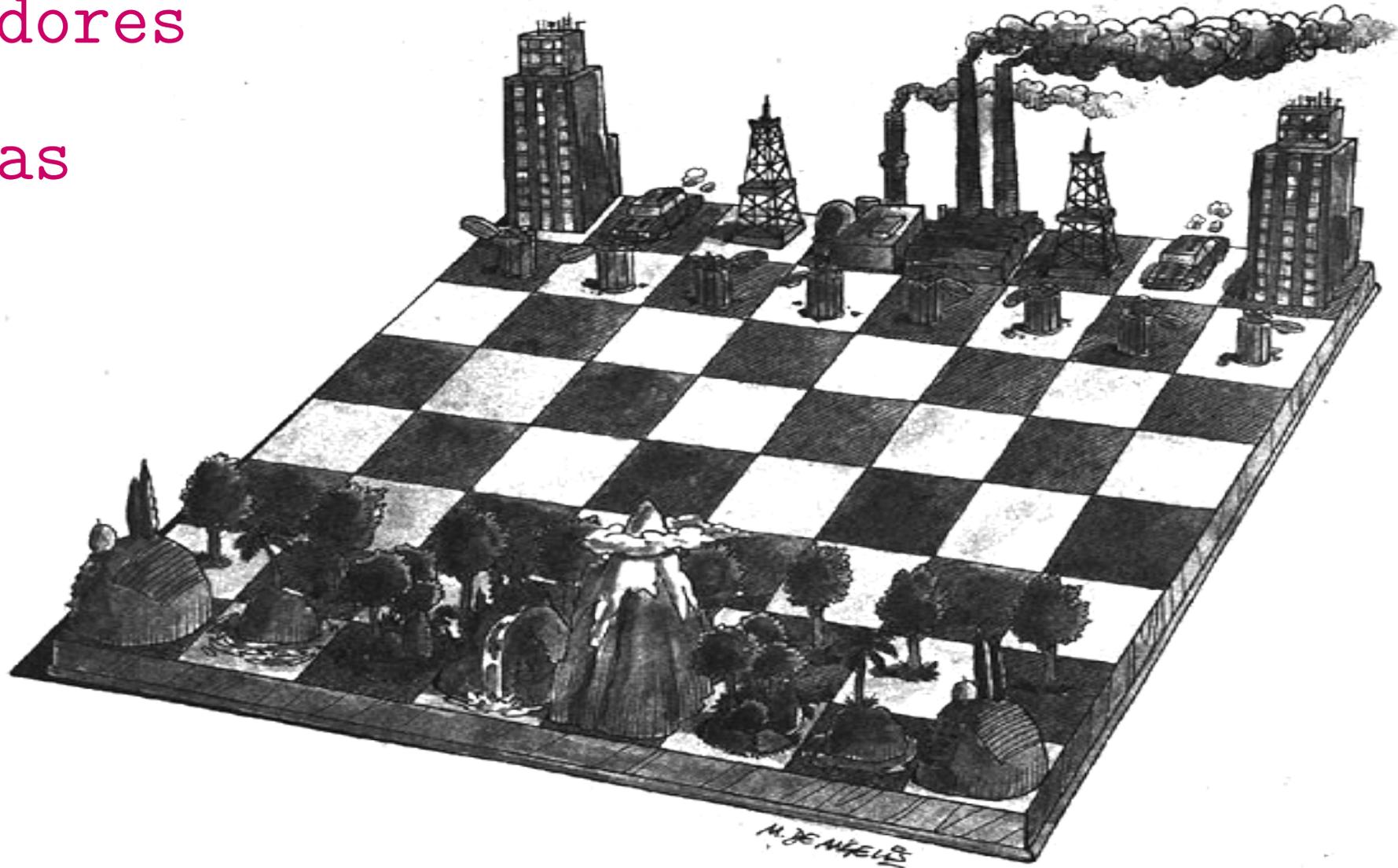
# Horizonte de tiempo y Paso de tiempo



- **Horizonte de Tiempo:** ventana de tiempo en la que vamos a realizar el estudio del sistema.
- **Paso de Tiempo:** intervalo seleccionado para realizar la integración de las ecuaciones del Sistema.



- Generadores
- Demandas
- Nodos
- Arcos



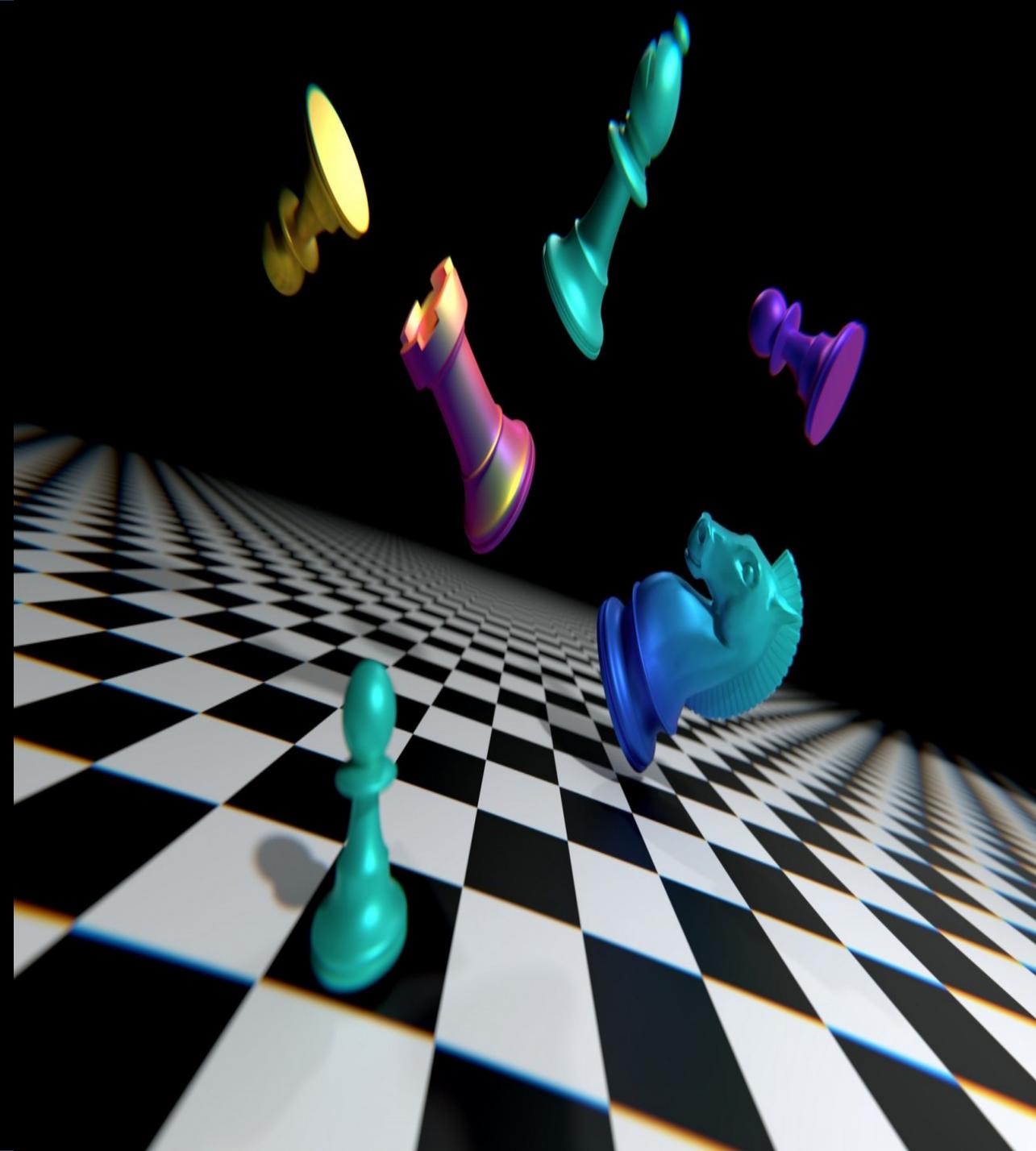
Implementación de SimSEE  
En cada paso de tiempo hay que:

Dimensionar Problema

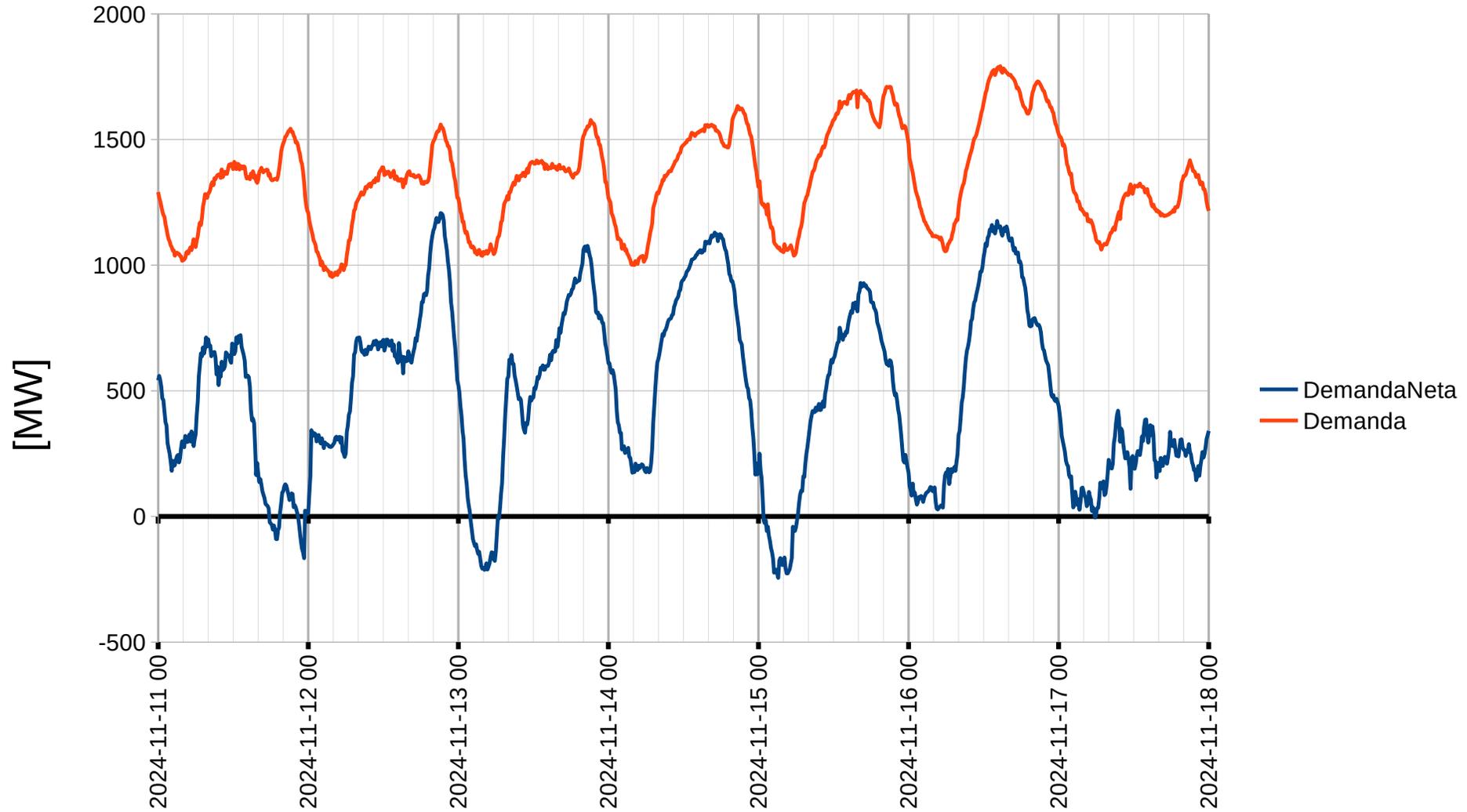
Armar Ecuaciones

Resolver Despacho

Leer Solución =  
Ejecutar Despacho



La incorporación masiva de Eólica y Solar lleva a la necesidad de considerar La Demanda Neta

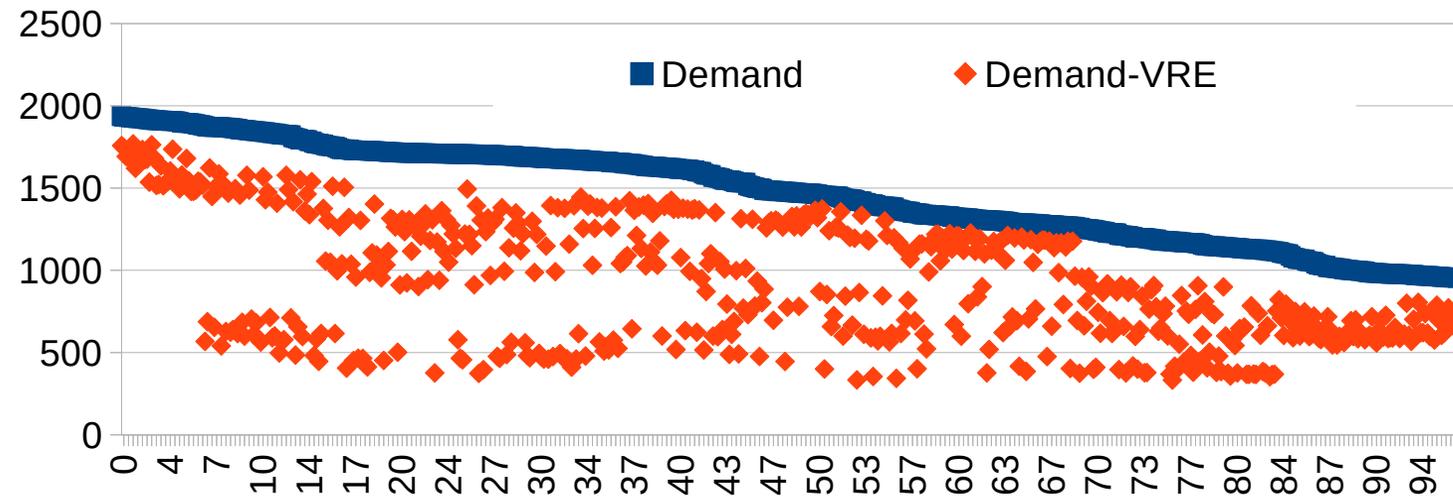
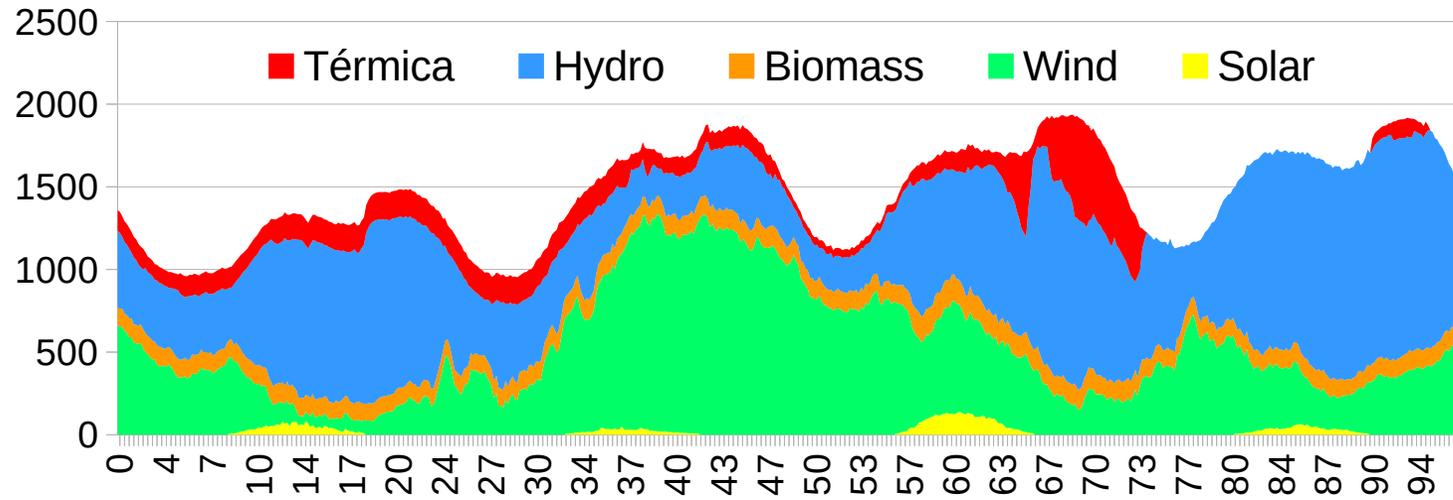


Construida con datos diezminutales SCADA <https://adme.com.uy>

# Resolución del despacho en un paso de tiempo en BLOQUES HORARIOS

(Postes en Uruguay, Patamares en Brasil)

- Para simular con pasos de tiempo superiores a la hora e intentar reflejar adecuadamente los requerimiento de balances de potencia, se suele recurrir a subdividir el paso de tiempo en BLOQUES HORARIOS re-ordendando las horas del paso de tiempo por su requerimiento de potencia.
- Los Bloques Horarios se llaman POSTES en SimSEE
- El uso de POSTES está en casi todas las herramientas de optimización/simulación de sistemas de energía desde hace décadas.
- Tradicionalmente se recurría a construir la curva Monótona de Carga (ordenando las horas de un paso de tiempo típico de mayor a menor potencia) y se seleccionaba el promedio de cada grupo.



Bloques horarios definidos a partir de la Monótona de Carga.

¿tiene sentido con sistemas con alta integración de eólica y solar?

solo un ejemplo, 4 days of july-2018-Uruguay

Source: ADME - SCADA ten-minute time series



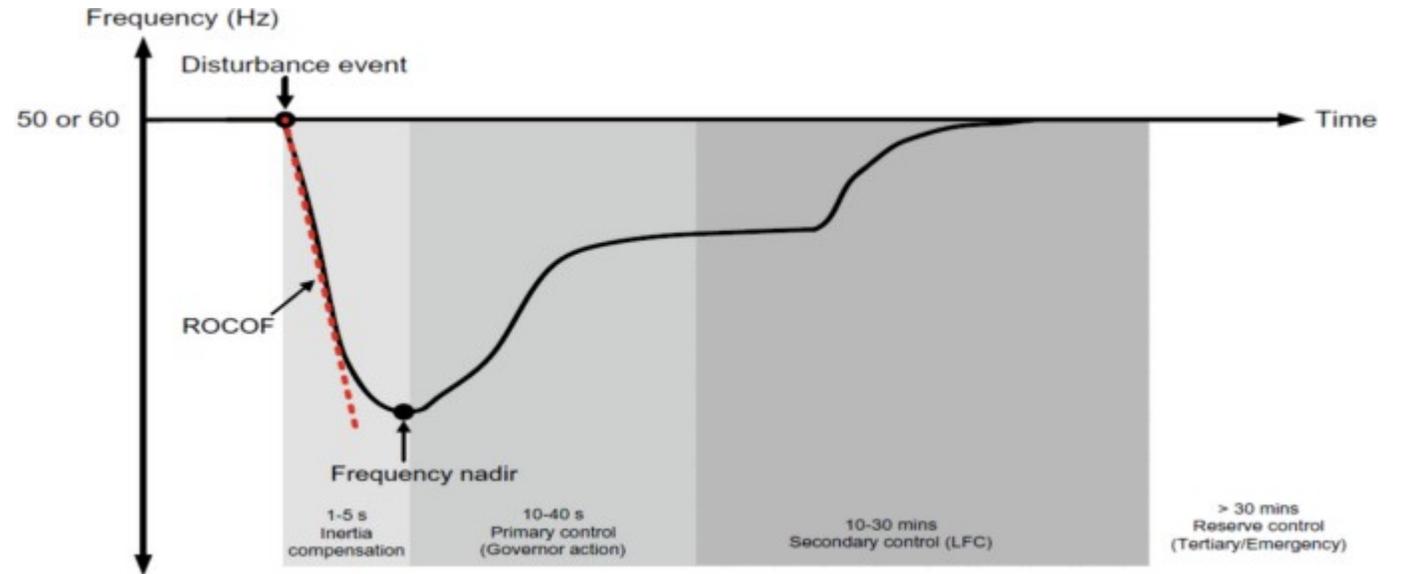
## SimSEE utiliza **Postizado-Dinámico**

En cada paso de tiempo, se crea la Demanda Neta con detalle horario sumando todas las demandas y restando las energías no-gestionables y en base a la monótona de la Demanda Neta se definen los POSTES para la resolución del paso.

# Riesgo de pérdida de inercia por aumento de ERNC

## Respuesta en Frecuencia

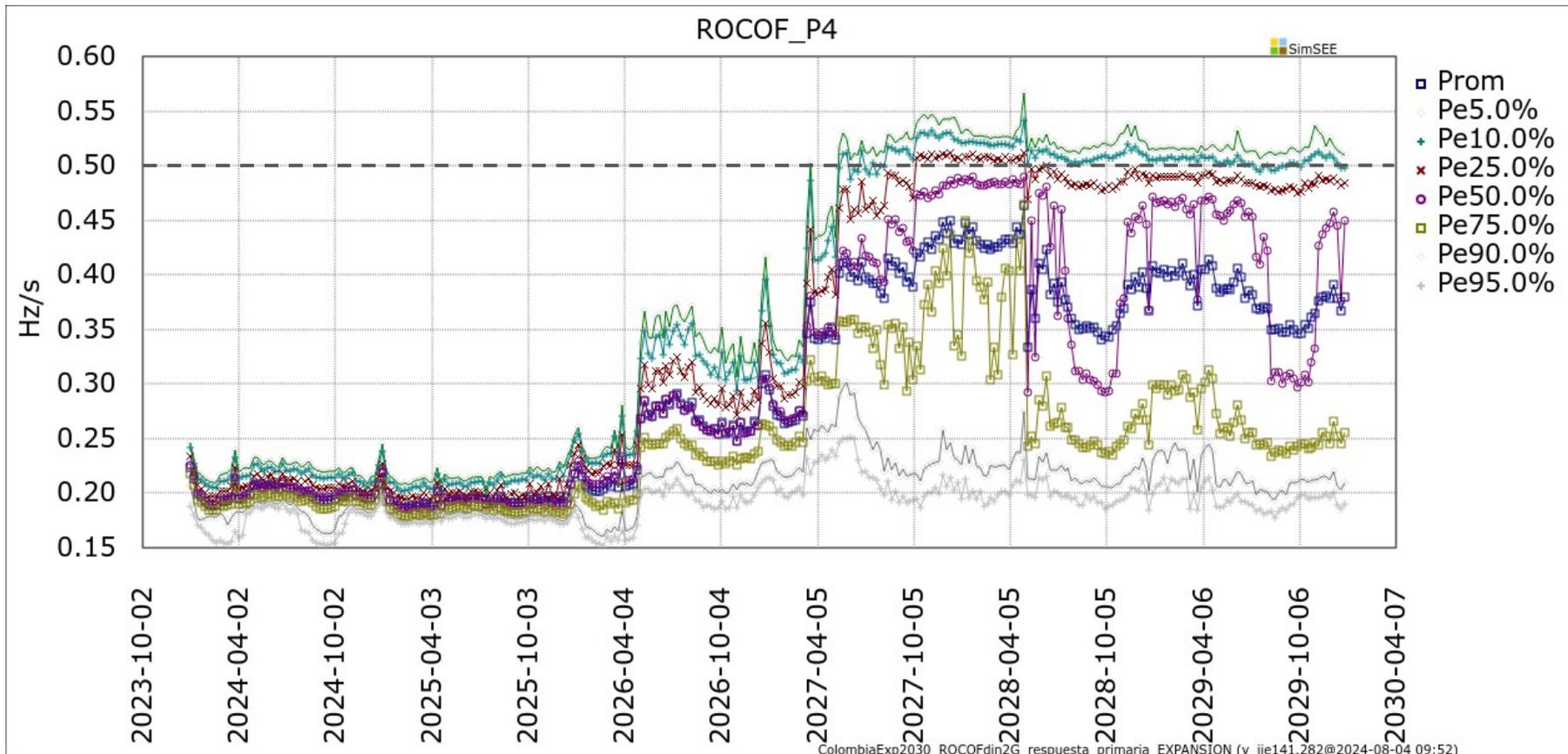
- Frecuencia límite inferior
- Tiempo de respuesta inicial ( $T_r$ ): 2s
- Tiempo de respuesta total ( $T_e$ ): 15s
- Límite de ROCOF propuesto por CREG:  
0.5 Hz/s



## Metodología

- 🏠 Fijación de demanda neta mínima
- 🔧 Incorporación al bucle de Optimización/Simulación del cumplimiento de ROCOF

# Casos de estudio y resultados



Evolución de ROCOF del SIN con plan de expansión base

# En etapa de pruebas:

## **SimDOR. Despacho Óptimo Regional**

- Demanda Neta por Región
- diferentes husos horarios

## **Restricciones de Red**

## **Restricciones de Corto-Circuito**

Operación óptima de un  
sistema dinámico



# Valor de un recurso almacenable



Comparación entre costo del presente y costo del futuro.

De no haber restricciones para el traslado en el tiempo, el costo marginal sería el mismo en todas las horas del futuro.

INCERTIDUMBRE DEL FUTURO.

MODELOS ESTOCASTICOS

PRONOSTICOS

# Estado de un Sistema Dinámico



- $X$  = Vector de información que capta todo lo relevante del pasado para calcular el futuro si se conocen las entradas de aquí en mas.

$$X(t) = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

# El despacho óptimo es un "problema de programación dinámica estocástica"

El uso de recursos almacenados (agua) hoy reduce los costos operativos del presente pero aumenta los del futuro, y viceversa.

Una política óptima es aquella que reduce el valor esperado del costo operativo futuro del sistema.

$$\text{Min } \langle CF \rangle$$

En el nivel óptimo, las variaciones causadas por las acciones de control sobre el presente equilibran las causadas sobre el futuro están en equilibrio.



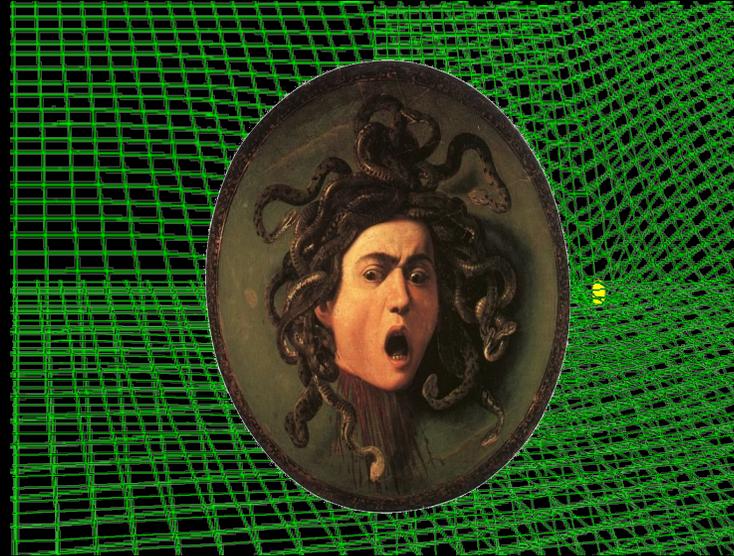
- *Dynamic Programming* 1957  
**Bellman recursion**

$$CF(\mathbf{X}, k) = \left\langle \min_{u_k} \left\{ ce(\mathbf{X}, u_k, r_k, k) + \rho CF(\mathbf{X}_{k+1}, k+1) \right\} \right\rangle_{\{r_k, r_{k+1}, \dots\}}$$

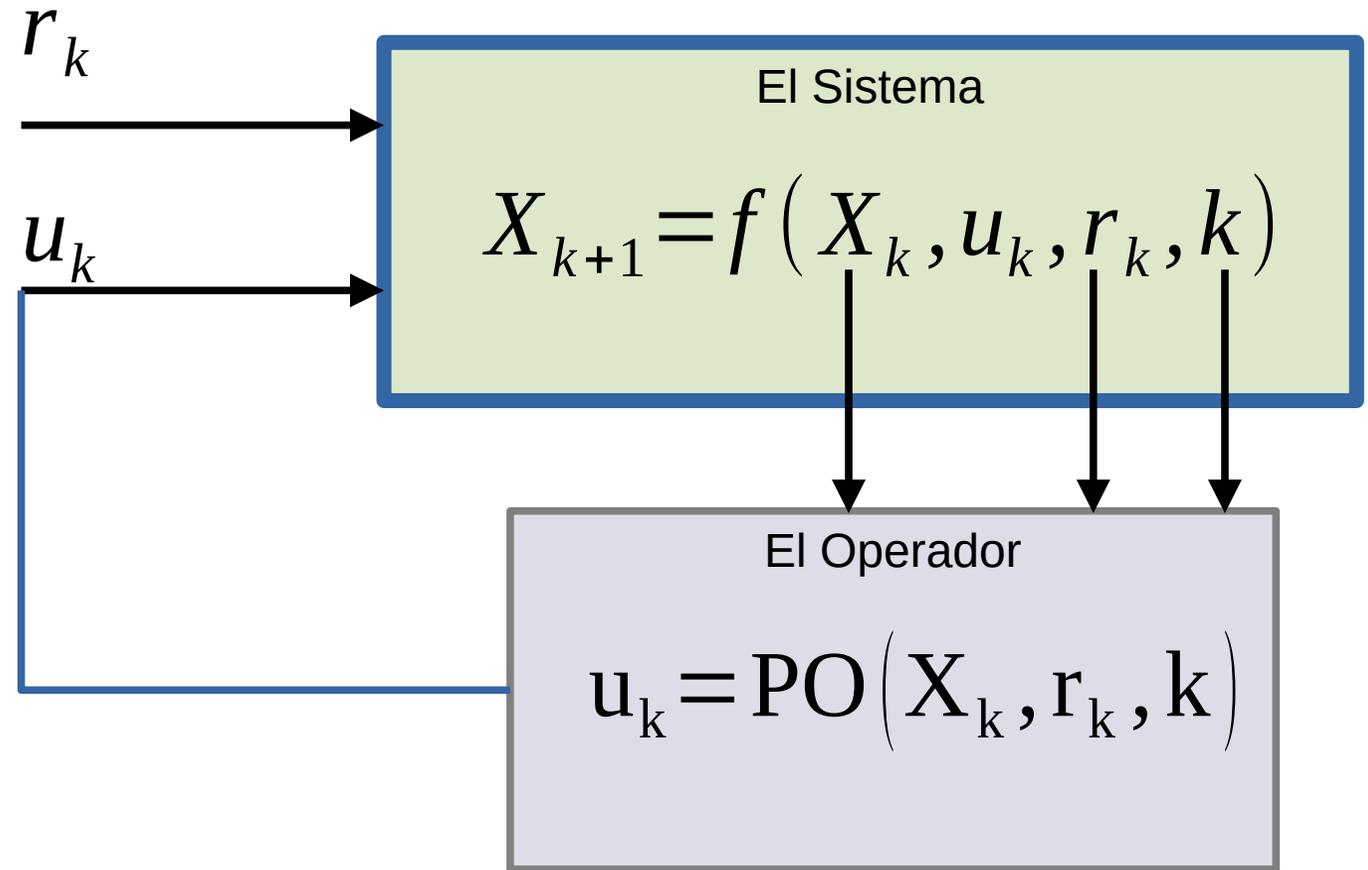


Richard Ernest Bellman (1920–1984)

La Maldición de la dimensionalidad de Bellman



# El Operador y la Política de Operación



## Las decisiones se vinculan temporalmente

El uso de los recursos almacenados (agua) hoy reduce los costos operativos del **presente** pero aumenta los costos operativos del **futuro**; y viceversa.

Una Política Óptima reduce el valor esperado del costo de operación futura del sistema.

Una Política Óptima equilibra los efectos de las acciones sobre los costos futuros y presentes.

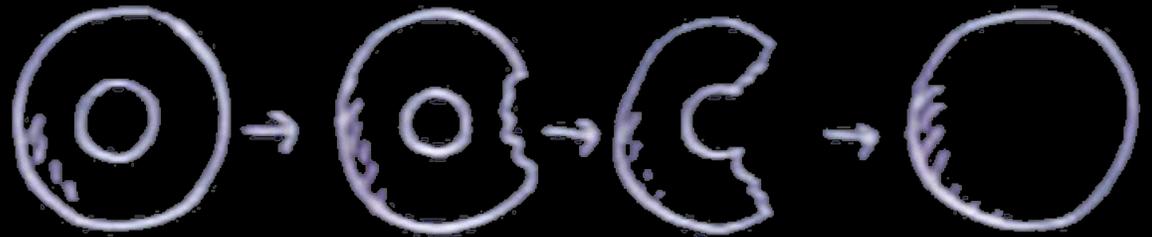


# El Costo Futuro

$$CF(X, k) = \left\langle \sum_{h=k}^{\infty} q^{h-k} C_h \right\rangle_{R, U}$$

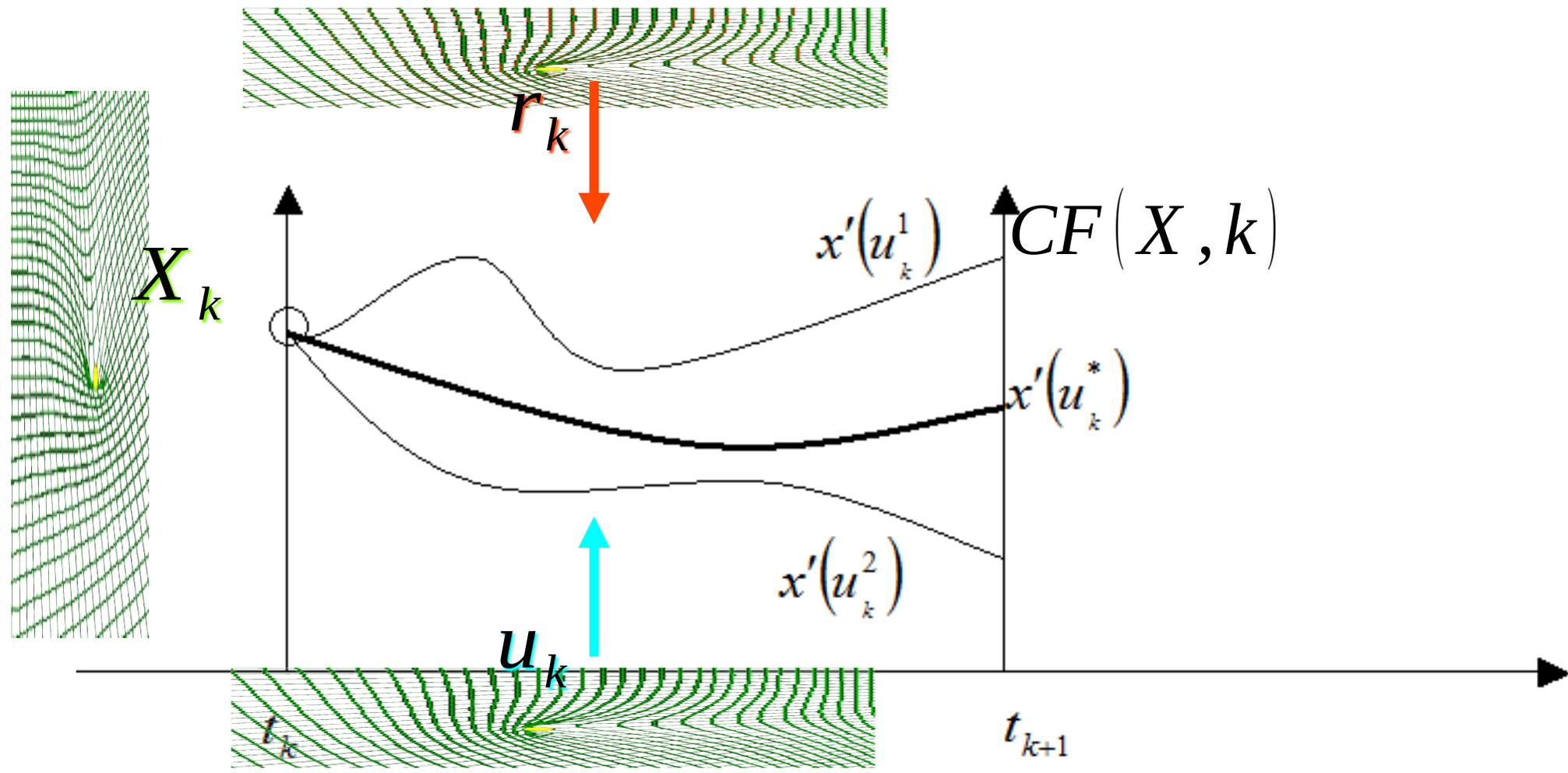
## Política Óptima

$$\min_{OP} \left\langle CF(X, k) \right\rangle_R$$





# La Maldición de la Dimensionalidad de Bellman



$$Dim(u) \times N_{X_1} \times N_{X_2} \dots \times N_{X_{Dim(X)}} \times N_{r_1} \times N_{r_2} \dots \times N_{r_{Dim(r)}} \times N_t$$

# Agregaciones Clásicas



- Encadenamiento de optimizaciones con diferente horizonte y paso temporal (meses, semanas, días, horas).  
Largo\_Plazo, Mediano\_Plazo y Corto\_Plazo.  
En el Largo plazo se consideran como variables de estado las PESADAS y se van agregando variables en el Mediano y en el Corto Plazo.
- En SimSEE con los CEGHs, posibilidad de reducción del estado de los procesos estocásticos.
- Subdivisión del paso de tiempo en POSTES, Bandas Horarias o PATAMARES.



# Modelado de la incertidumbre ( Procesos Estocásticos )



## Fuentes de aleatoriedad

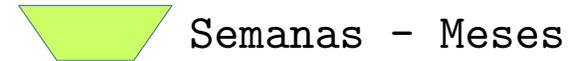
- Demanda
- Caudales de aportes hídricos
- Velocidad del viento
- Radiación solar
- Precio de los mercados vecinos
- Precios de los combustibles
- Disponibilidad de combustibles
- Roturas fortuitas

# Caracterización de la variabilidad

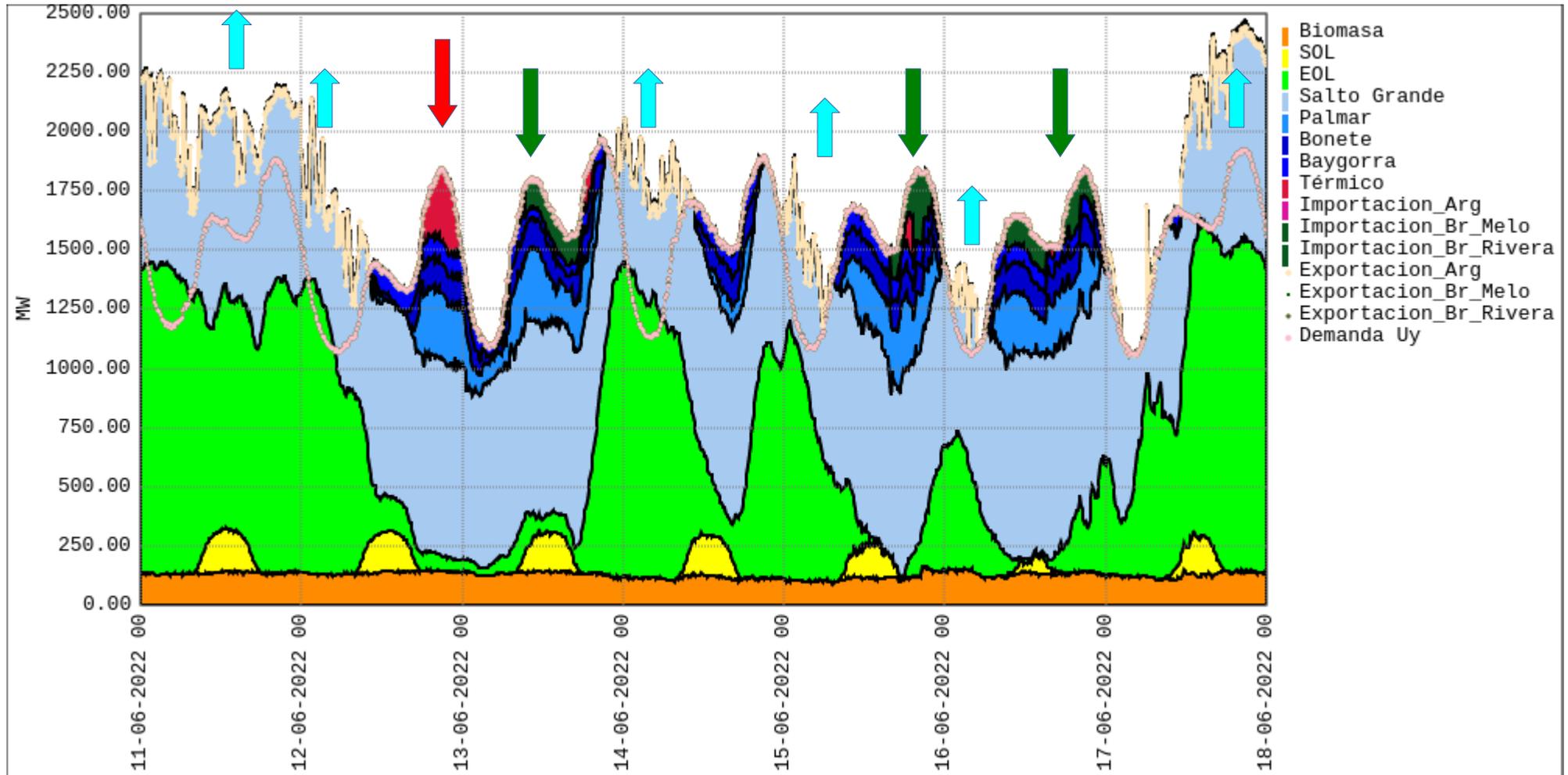
Caudales hidráulicos



Solar y eólica



# Ejemplo: Importancia de los INTERCAMBIOS OCASIONALES



# Modelos CEGH

- Reproduce los histogramas de amplitudes de las series.
- Reproduce las correlaciones entre las series y con sus pasados..

*Espacio Gaussiano:  
Sistema lineal multi-variable  
alimentado por vector de  
Ruido Blanco Gaussiano.*

$$X_{k+1} = \sum_{h=0}^{h=n-1} A_h X_{k-h} + \sum_{h=0}^{h=m-1} B_h R_{k-h}$$

Pronósticos

NLT

NLT

NLT

NLT

NLT

NLT

*Espacio Real*

# Operador Sin Pronósticos.



# Operador Con Pronósticos.



# Facilidad de integración de PRONÓSTICO

$$X_{k+1} = \sum_{h=0}^{h=n_r-1} A_h X_{k-h} + S_k + F_k \sum_{h=0}^{h=m-1} B_h R_{k-h}$$

Los **s** sesgos (**s**) cambian la Guía del probabilidad 50% y los factores de atenuación (**f**) regulan la inyección de ruido, permitiendo ir desde un Pronóstico Determinístico (ruido nulo) a la desaparición del pronóstico (ruido histórico).

$$S_k = \begin{bmatrix} S_{1,k} \\ \dots \\ S_{n,k} \end{bmatrix}$$

$$F_k = \begin{bmatrix} f_{1,k} & 0 & \dots & 0 \\ 0 & f_{2,k} & \dots & 0 \\ 0 & \dots & 0 & f_{n,k} \end{bmatrix}$$

# Estado del proceso estocástico

$$X_{k+1} = \sum_{h=0}^{n_r-1} A_h X_{k-h} + \sum_{h=0}^{m-1} B_h R_{k-h}$$

La información del pasado, necesaria para calcular la salida del sistema lineal en el espacio gaussiano, pasa a formar parte de El Estado del Sistema.

$$X_k, X_{k-1}, \dots, X_{k-(n_r-1)}$$

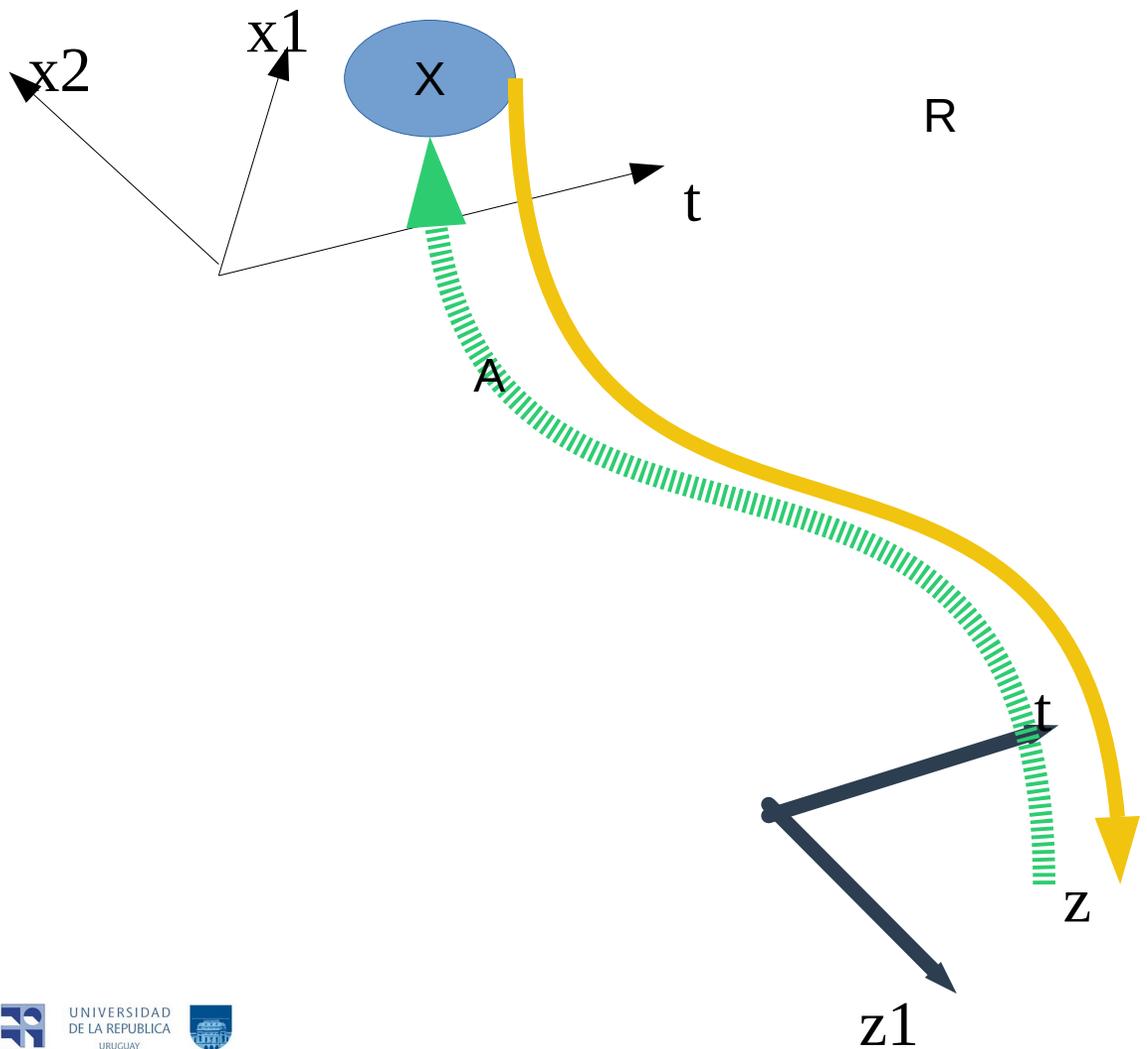
# Maldición de Bellman.

$$u_k = P(x_k, r_k, k)$$



$$\text{Dim}(u) \times \text{Dim}(x) \times \text{Dim}(r) \times \text{Dim}(k)$$

# Reducción del Espacio de Estado en modelos CEGH



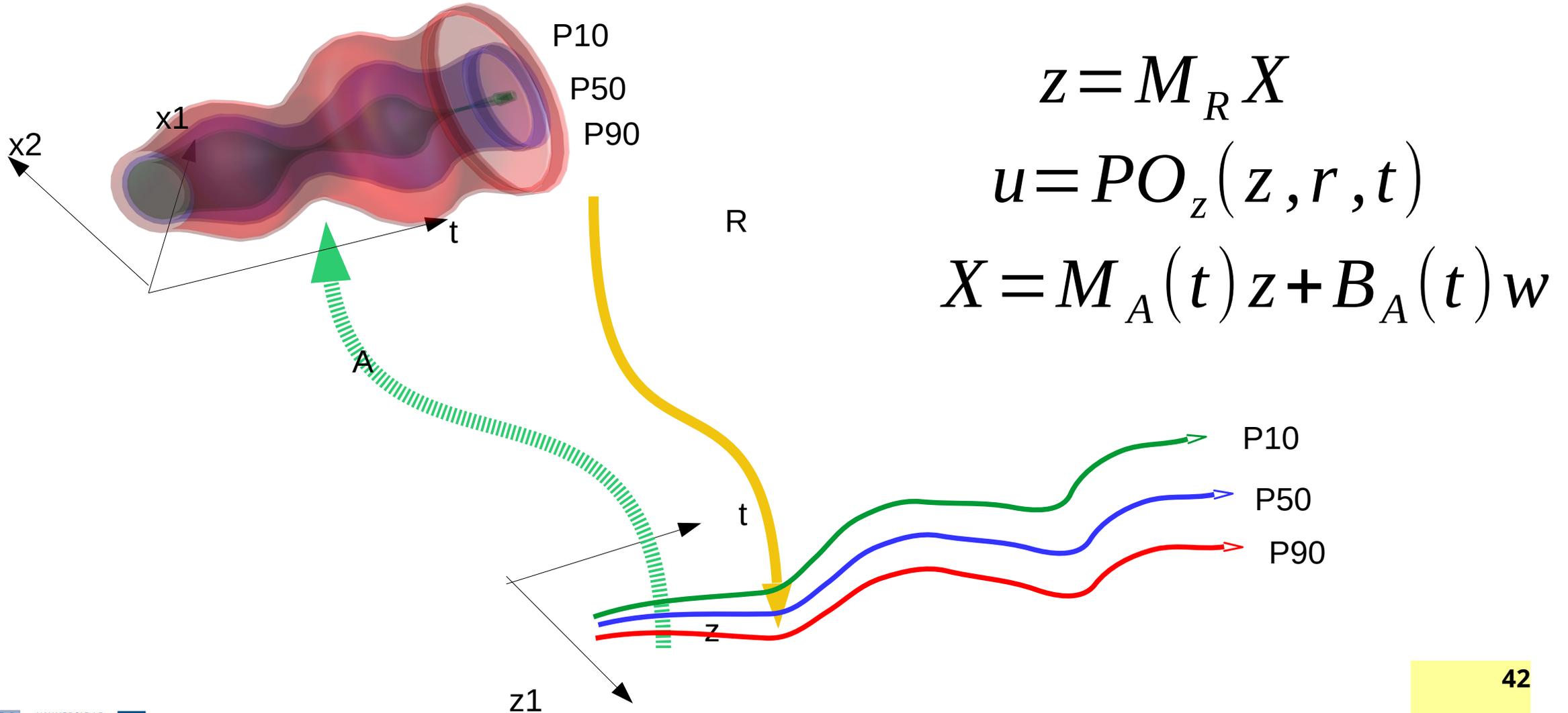
$$z = R(X)$$

$$u = PO_z(z, r, t)$$

$$X = A(z, w)$$

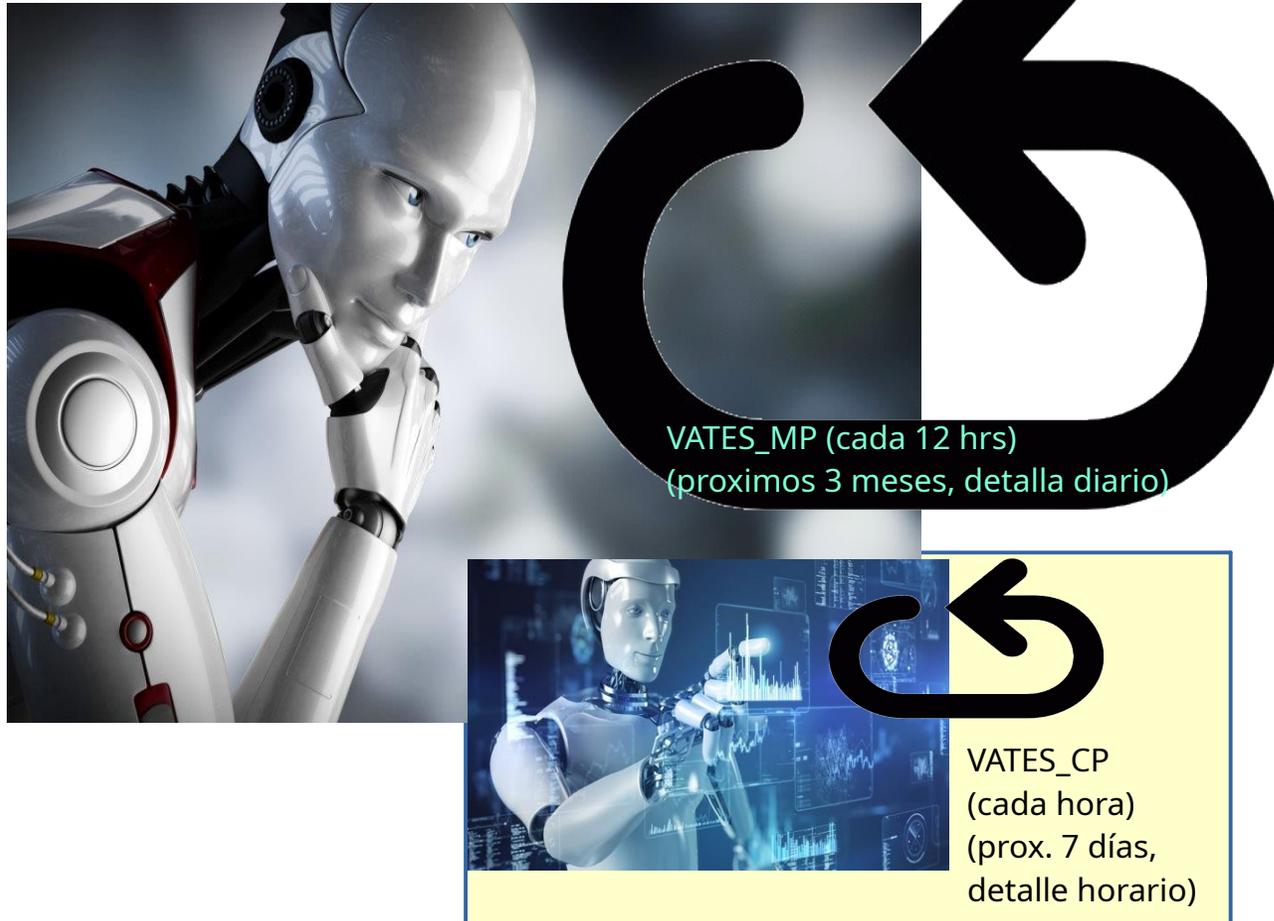
W es el ruido que permite poblar el volumen de X que mapea en el z dado.

# Tratamiento del pronóstico en el espacio Gaussiano con reducción



# Administración del Mercado Eléctrico

Robots VATES: Despacho de energía con asimilación del estado del SIN y pronósticos de forma continua.



VATES\_MP (cada 12 hrs)  
 (proximos 3 meses, detalla diario)

VATES\_CP  
 (cada hora)  
 (prox. 7 días,  
 detalle horario)

- En ADME contamos con dos Robots que están permanentemente resolviendo el despacho óptimo.
- Ambos asimilan la información del estado del sistema y de las previsiones de lluvia, viento, radiación solar y Demanda y resuelven la política óptima de operación.
- Un Robot analiza los próximos tres meses con detalle diario y publica los resultados dos veces al día.
- El otro analiza los próximos siete días con detalle horario y publica los resultados cada hora.
- Ambos robots utilizan la Recursión Bellman, lo que nos condena a no poder seguir añadiendo variables de estado y detalles al modelo del sistema.
- Esto nos llevó a desarrollar una nueva generación de Robots basados en técnicas de Inteligencia Artificial para intentar escapar de la Maldición Bellman.

PARTE 2)

# El Tractorcito

Aprendizaje por refuerzo

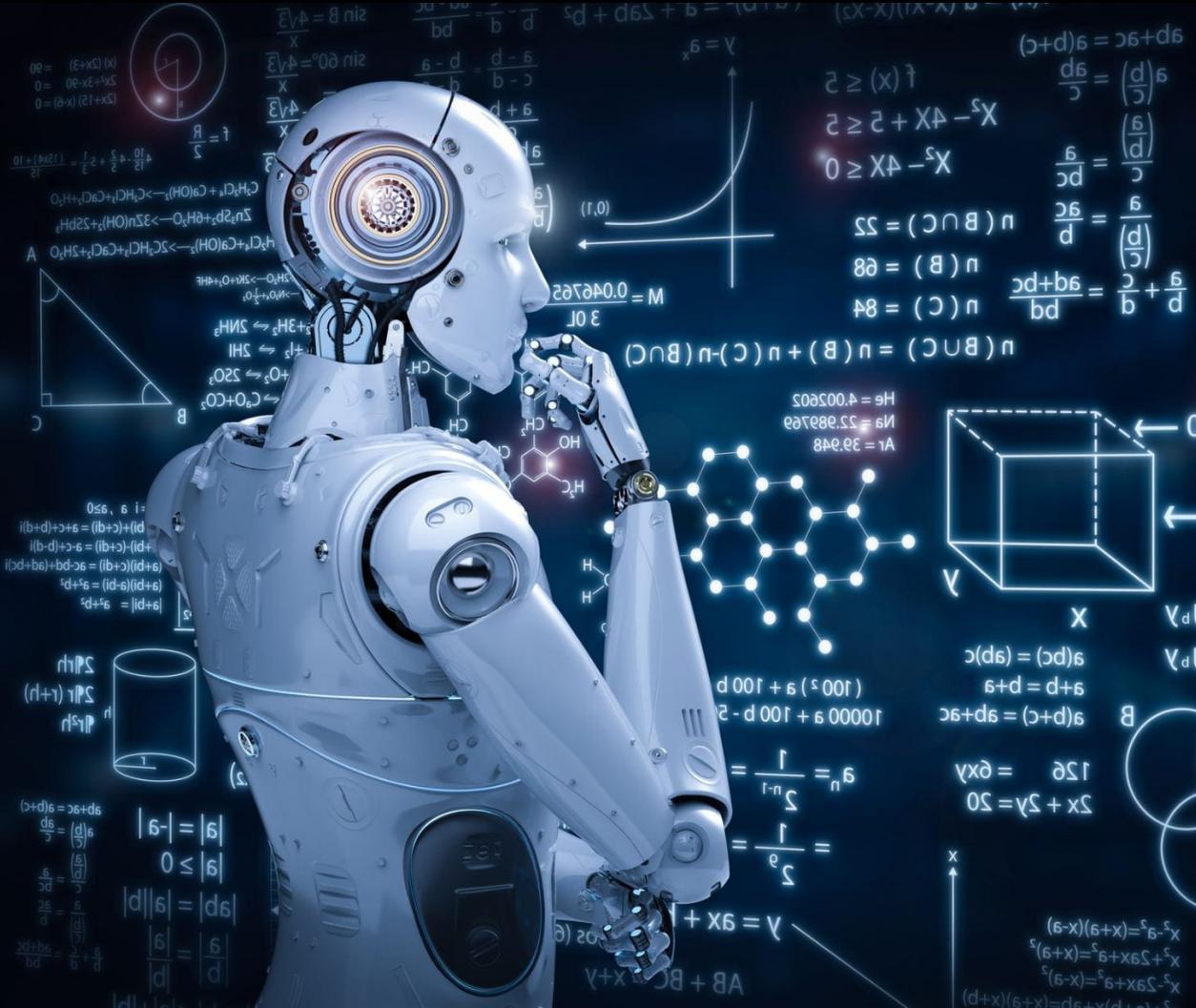
- + Control de Varianza
- + Control de exploración
- + Regularizaciones Espaciales
- + Parsimonia
- + Aprendizaje CONTINUO





Aprendiendo por ensayo y error a operar el sistema de energía eléctrica

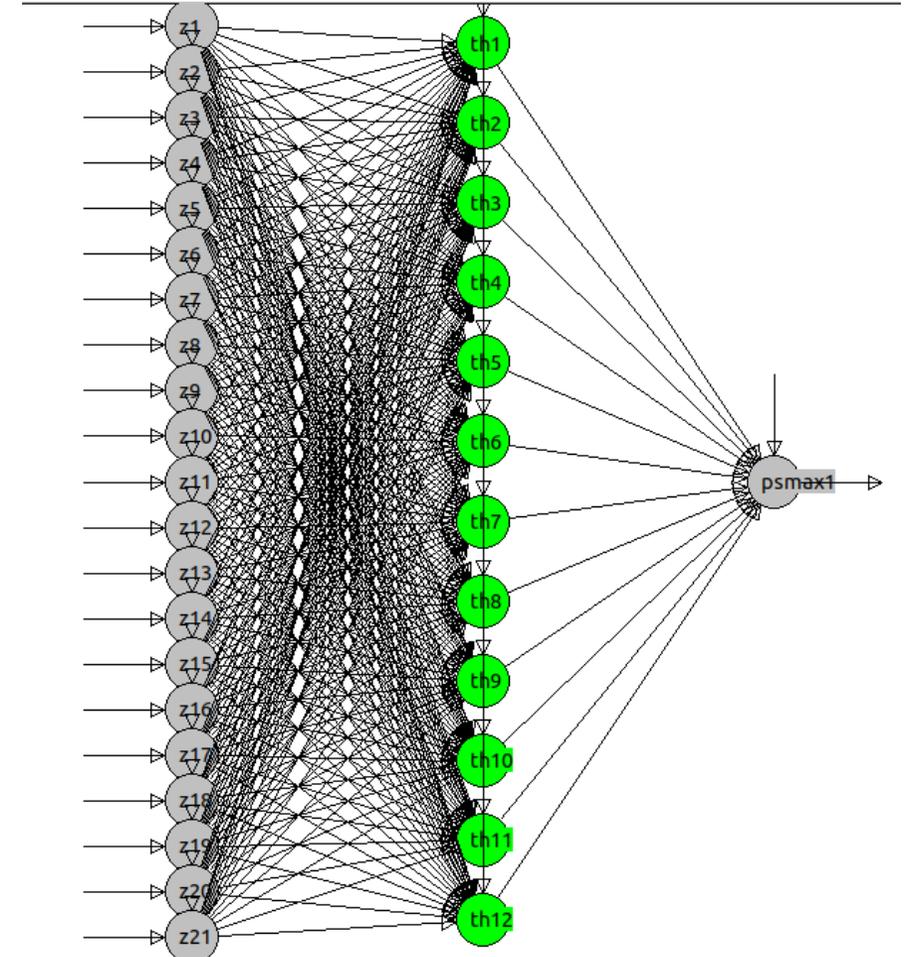
# Aprendizaje Automático ( machine learning )



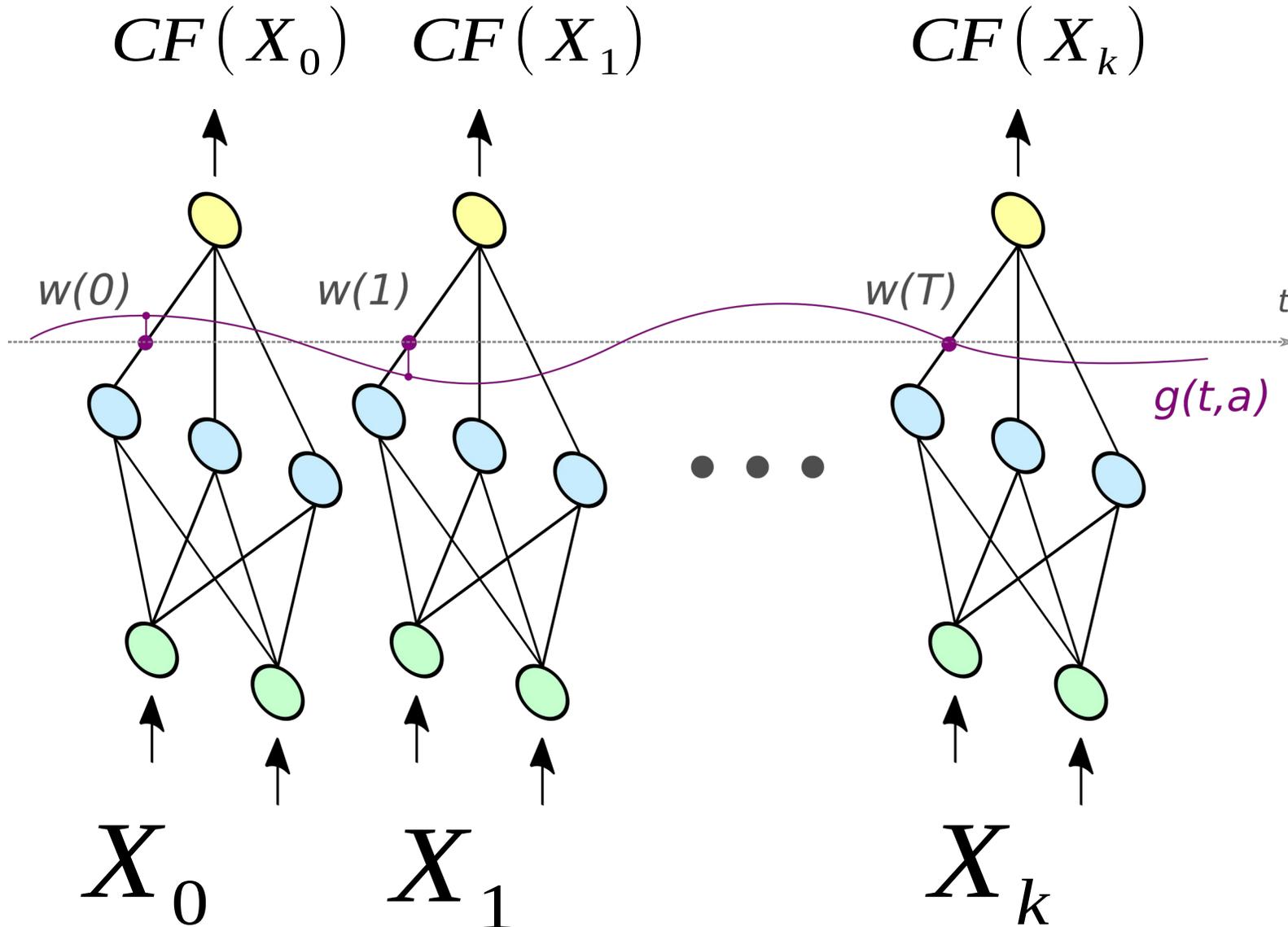
$$\min_u \left\{ ce(X_k, u_k, r_k, k) + CF(X_s, k+1) \right\}$$
$$@ \begin{cases} u \in \Omega(X_k, r_k, k) \\ X_s = f(X_k, u_k, r_k, k) \end{cases}$$

# Programación dinámica estocástica aproximada

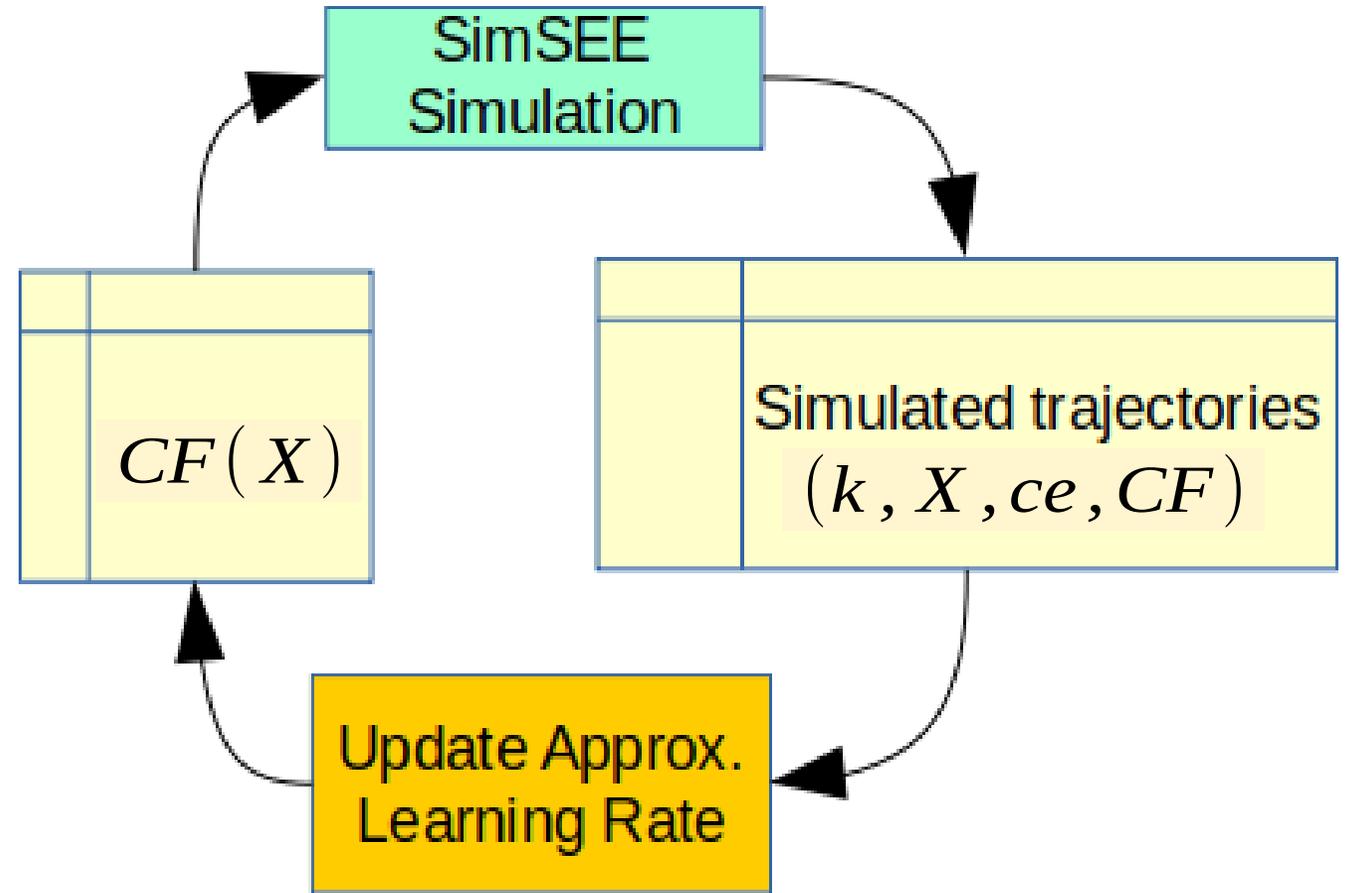
*... dada una aproximación de  $CF(X)$  puedo simular*



Conjunto de redes neuronales, una por paso de tiempo



# Bucle de aprendizaje



# Trayectorias (nx x nw):

Estados iniciales

Semillas Aleatorias

$$\left\{ X_0^1, X_0^2, \dots, X_0^{nx} \right\} \times \left\{ W_1, W_2, \dots, W_{nw} \right\}$$



Trayectorias

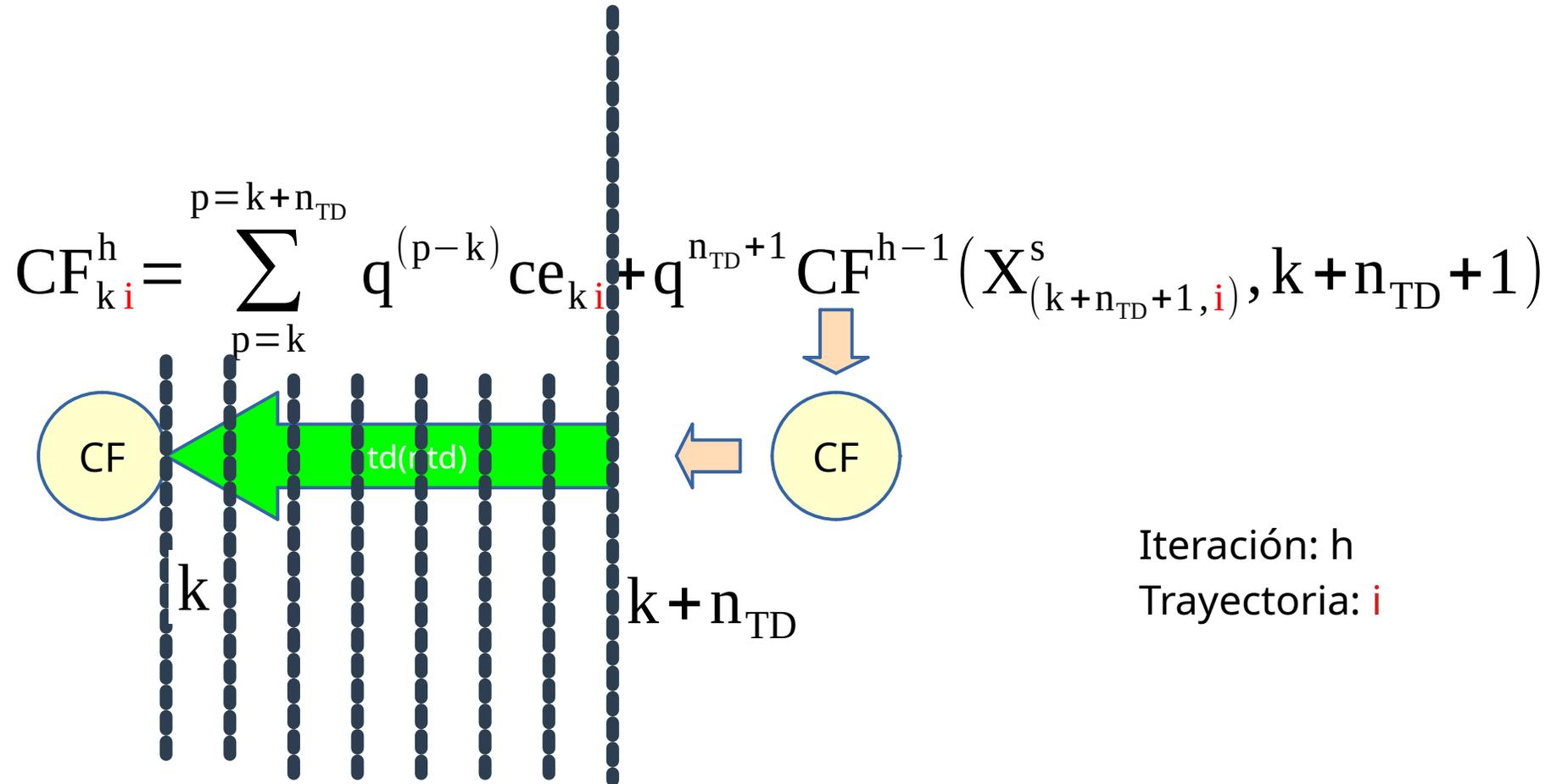
$$\left\{ \begin{array}{l} (X_0^1, w_1), (X_0^1, w_2), \dots, (X_0^1, w_{nw}) \\ (X_0^2, w_1), (X_0^2, w_2), \dots, (X_0^2, w_{nw}) \\ \dots \\ (X_0^{nx}, w_1), (X_0^{nx}, w_2), \dots, (X_0^{nx}, w_{nw}) \end{array} \right\}$$

$(X_0^1, w_1)$

k	X						MEE	ce	CF(X,k+1)
	x1	x2	x3	x4	x5	x6			
0	6339.434	560.223	1040.023	-1.378	-0.981	-0.811	7	-227978	-297041
1	6188.706	171.883	411.939	-1.368	-0.998	-0.846	7	-188945	-119353
2	6046.920	113.600	411.939	-1.339	-0.954	-0.880	7	-548	-17670
3	5929.287	113.600	411.939	-1.334	-1.021	-0.994	7	-206043	61942
4	5803.822	113.600	411.939	-1.333	-0.924	-1.113	7	-115979	126864
5	5687.223	113.600	411.939	-1.330	-0.960	-1.160	7	-58225	162629
6	5577.711	113.600	411.939	-1.325	-0.967	-1.128	7	-91917	196991
7	5474.921	113.600	411.939	-1.317	-0.963	-1.084	7	68926	231503
8	5378.466	113.600	383.063	-1.307	-1.115	-1.038	7	21433	255490
9	5286.142	112.375	410.533	-1.300	-1.109	-1.044	7	-142275	344703
10	5199.748	113.600	411.939	-1.296	-1.006	-1.079	7	-139002	416420
11	5119.304	113.600	411.939	-1.287	-0.884	-1.102	7	-44928	456890
12	5044.897	113.600	411.939	-1.212	-0.799	-1.105	7	-70140	486965

# Estimación de CF a partir de la información de las trayectorias

td = Time Difference



Iteración:  $h$   
 Trayectoria:  $i$

# Claves para el éxito de la implementación

Técnicas de reducción de varianza (Common Random Number )

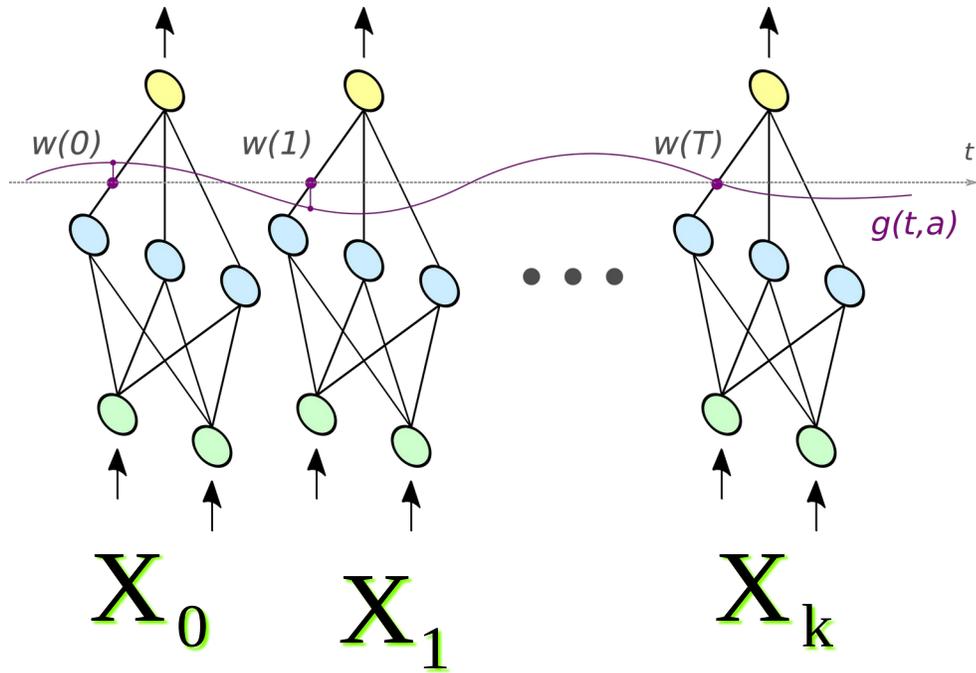
Evitar convergencia de las trayectorias durante la exploración.

+ mezcla de Time-Differences acorde con las constantes de tiempo del sistema. (Modos de Evolución)

Regularizaciones:

- + Red Neuronal de estructura simple
- + En los parámetros (Ridge o Lasso)
- + En la variación temporal de los parámetros.

# Representación de CF(X,k) y regularizaciones



$$L = \sum_{k,g} L_{kg} + \lambda \sum_k \|\theta_k\|^p + \beta \sum_{k=2} \|\theta_k - \theta_{k-1}\|^2$$

- Parsimony(t)
  - TD
  - Power series and harmonics

- p=2 : Ridge
- p=1 : Lasso

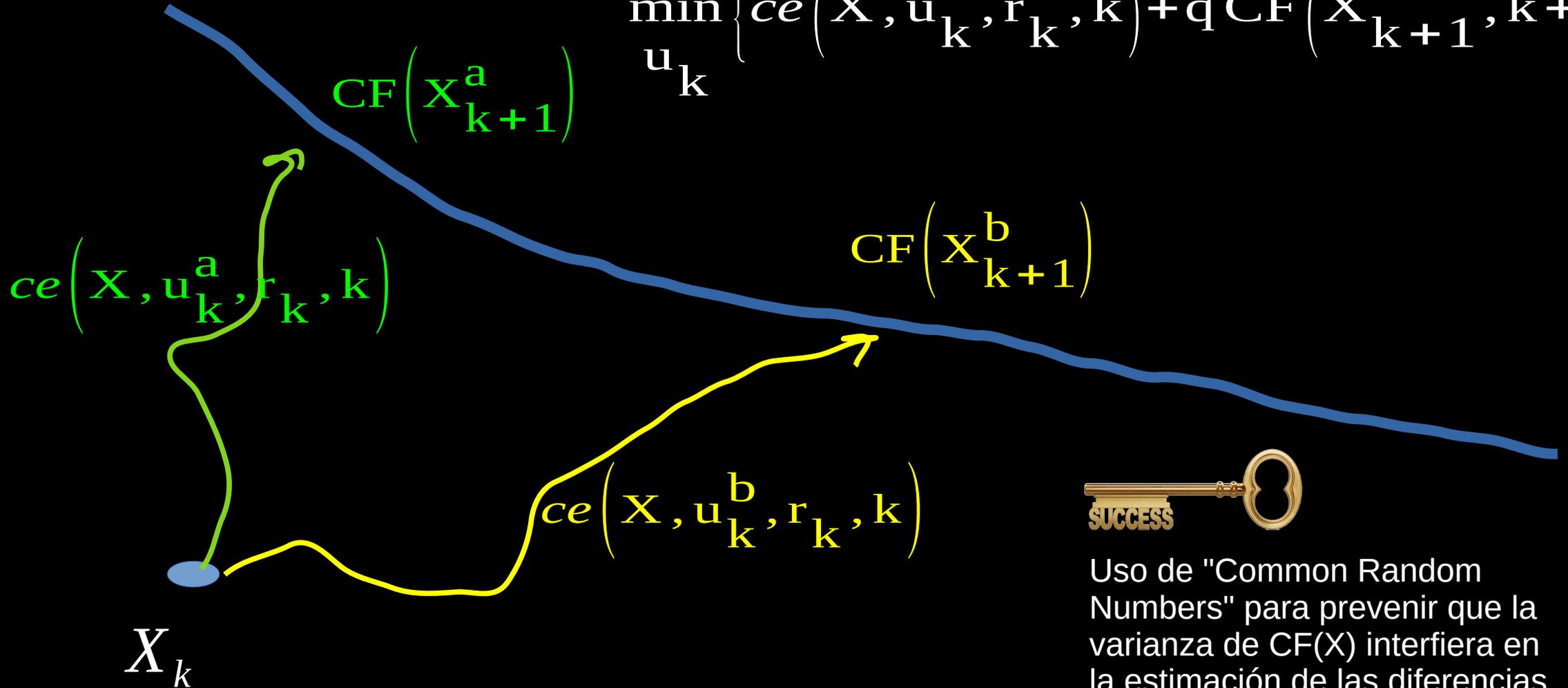


Approximate the differences instead of the function itself

$$L_{kg} = \frac{1}{4N^2} \sum_{i \neq j \in g} ((M(X_{kj}, \theta_k) - M(X_{ki}, \theta_k)) - (FC_{kj} - FC_{ki}))^2$$

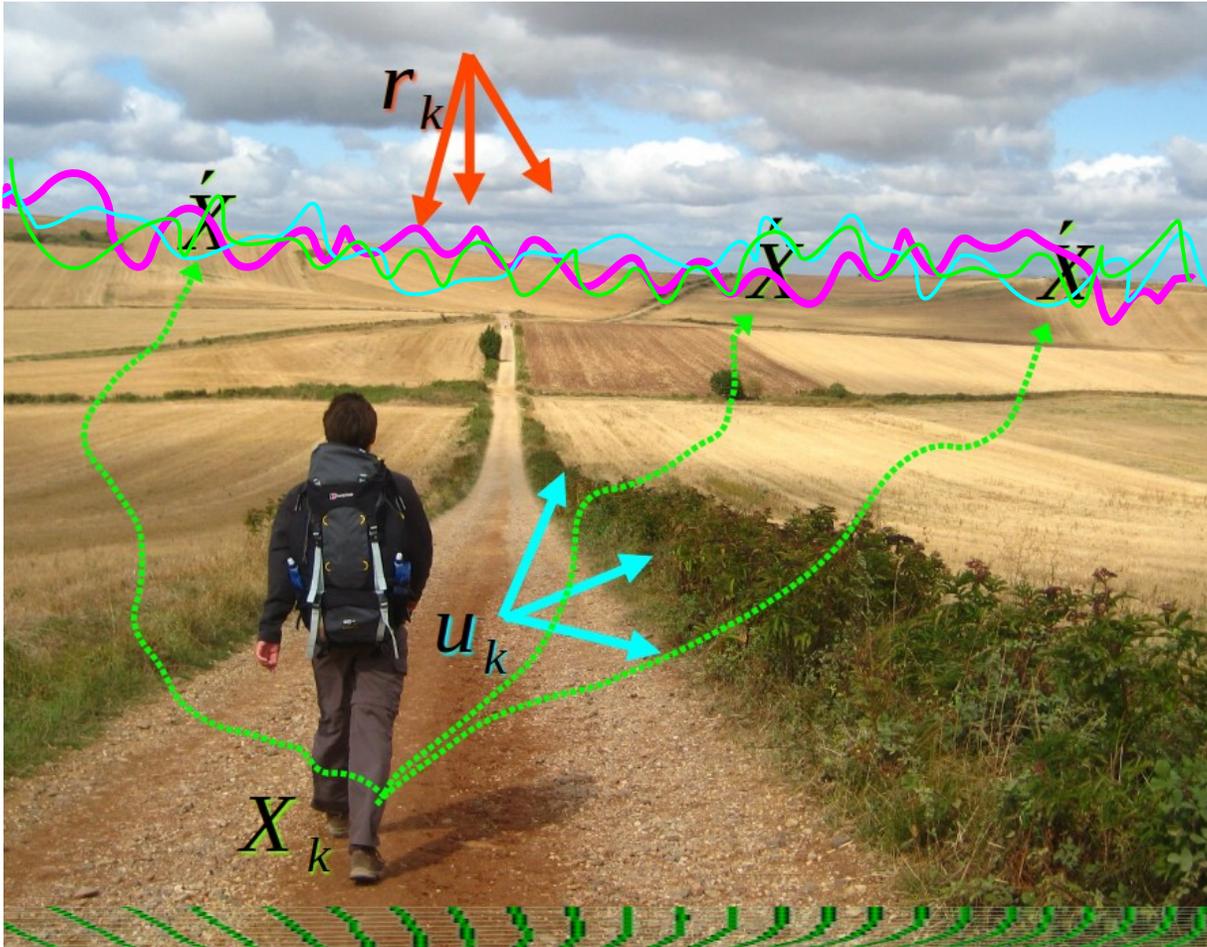
# La información está en las diferencias de $CF(X,k)$ respecto de $X$

$$\min_{u_k} \left\{ ce(X, u_k, r_k, k) + q CF(X_{k+1}, k+1) \right\}$$

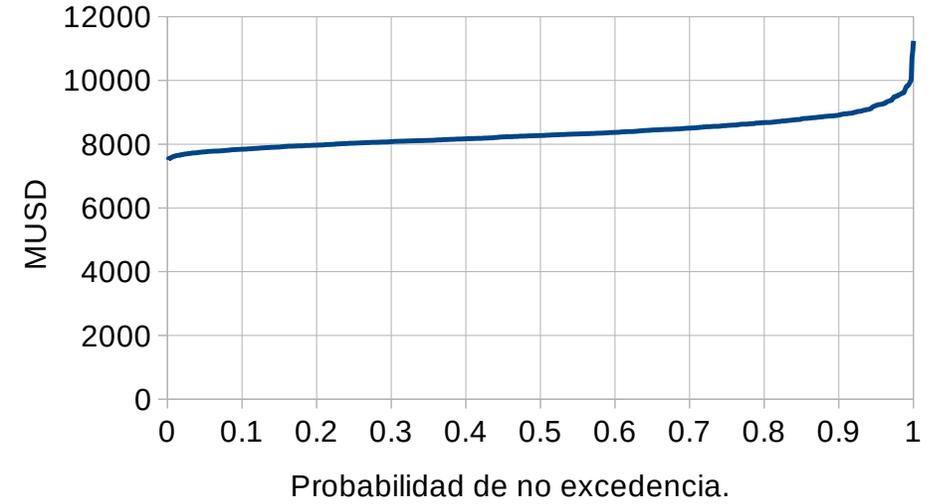


Uso de "Common Random Numbers" para prevenir que la varianza de  $CF(X)$  interfiera en la estimación de las diferencias :  $CF(X_a) - CF(X_b)$  en el estado de llegada

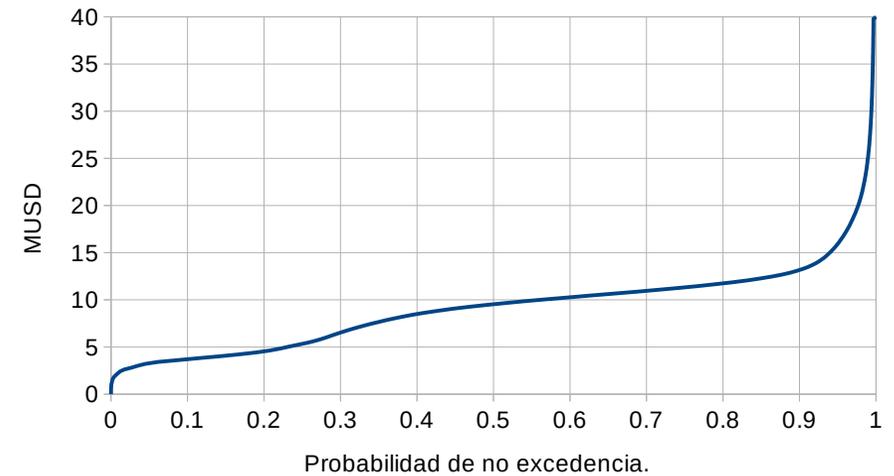
Por qué son necesarias las técnicas de reducción de la varianza.  
 Simulación 1000 crónicas con paso de tiempo semanal de Uruguay.



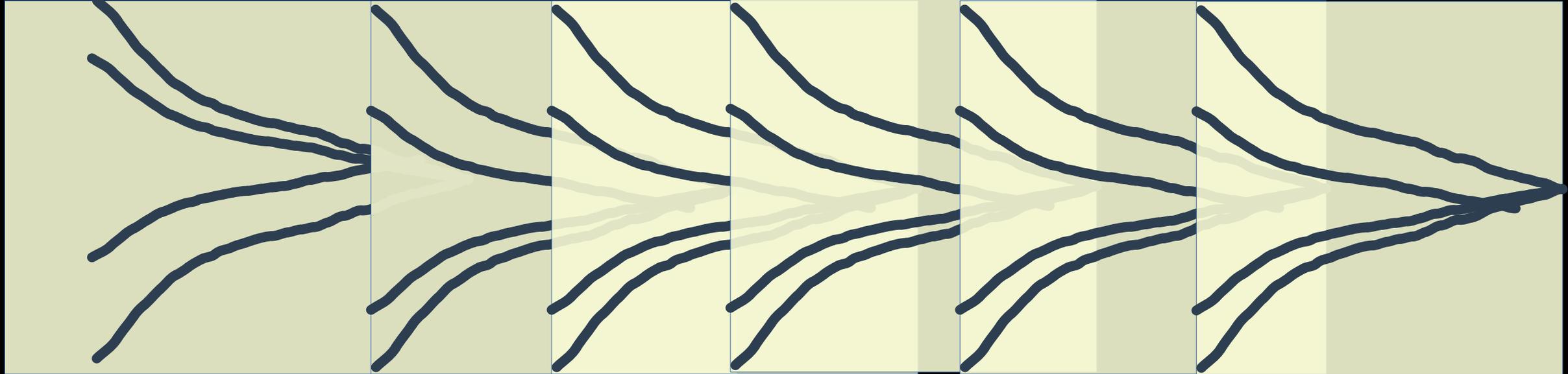
Distribución CF(X) - Uruguay



Distribución del costo de etapa (semanal).



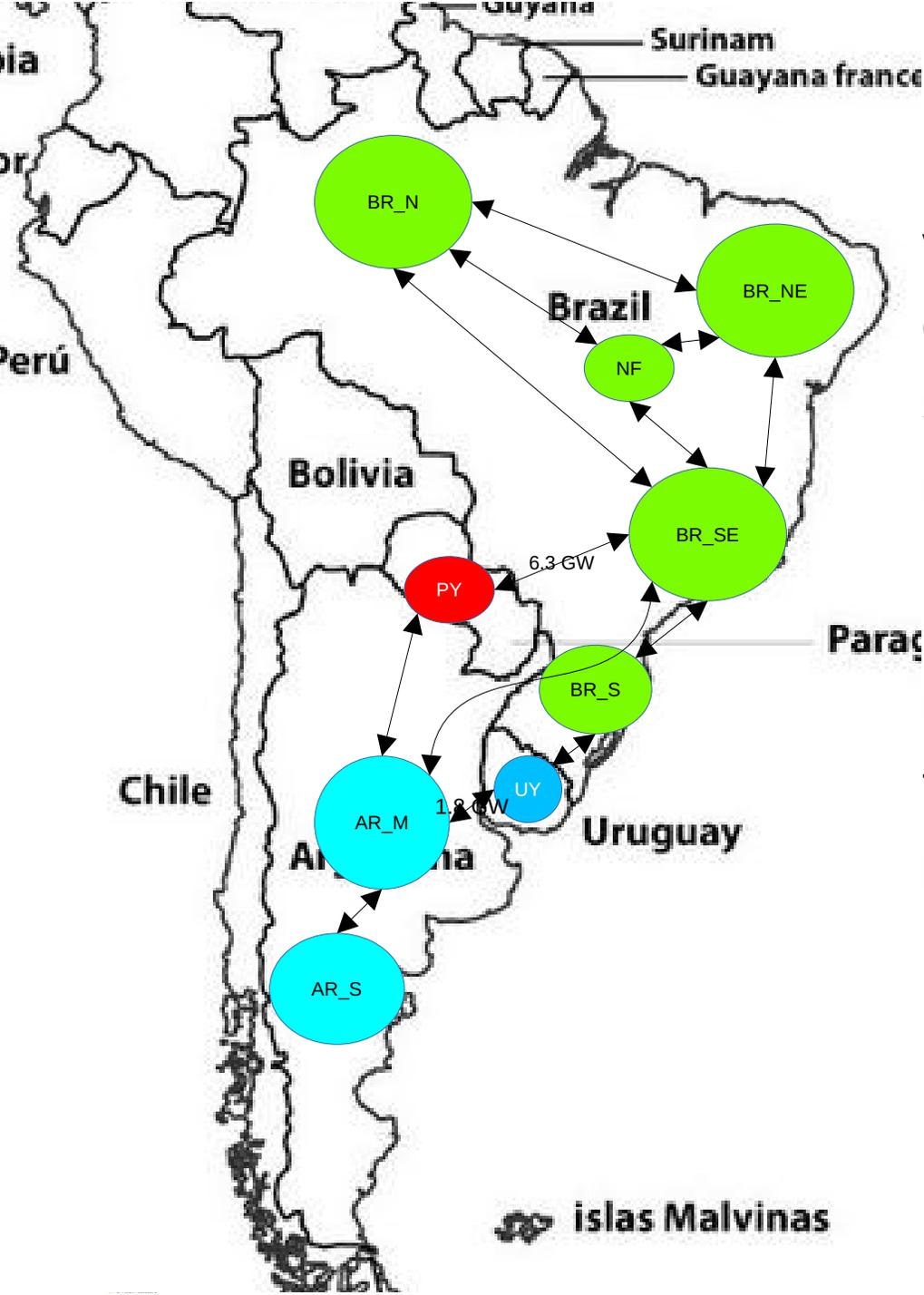
# Estrategias de exploración



En un sistema estable, el uso de CRN conduce a la convergencia de trayectorias y a la consiguiente pérdida de capacidad de exploración. Cada estado de los pasos de NTD se aprovecha para recuperar dicha capacidad.



En un sistema de funcionamiento continuo, el aprendizaje también se procesa de forma continua. Con el transcurso de cada hora real, el vector de neuronas se va desplazando, descartando el array de neuronas correspondiente a la hora transcurrida y repitiendo la última para continuar en el bucle de aprendizaje.



# Simulación UY+BR+PY+AG

variables de estado: 76

discretización del espacio de estado:

$$52 * 3 * 5^{76} = 2 \times 10^{55}$$

Recursión de Bellman memoria:

$$2 * 10^{55} * 10 / 1E9 = 2 * 10^{47} \text{GB}$$

Dimensión Simplex sobre: 5000x2000

Tiempo de resolución Simplex: 64 ms

hora de resolver la recursión de Bellman:

$$2 * 10^{55} * 64 / 1000 / 3600 / 24 / 365 / 1000 = 4 * 10^{43} \text{ milenios!!!!}$$

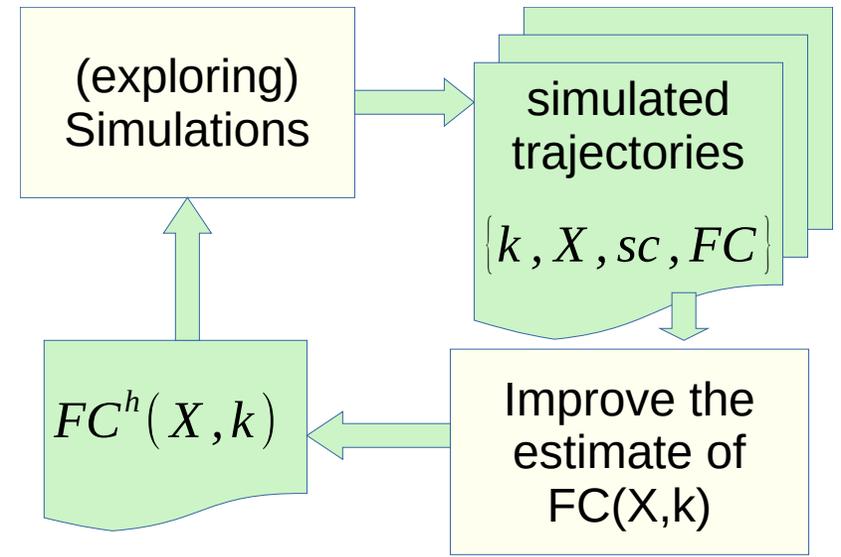


# Administración del Mercado Eléctrico

## Robots

### VatesIA\_MP y VatesIA\_CP

Segunda generación con inteligencia artificial



PARTE 3)

# OddFace+HC

Optimizador genético con Herencia Cultural  
Optimización de Inversiones

+ Optimización Genética  
+ Herencia Cultural



# OddFace

Optimizador distribuido de Funciones de alto costo de evaluación

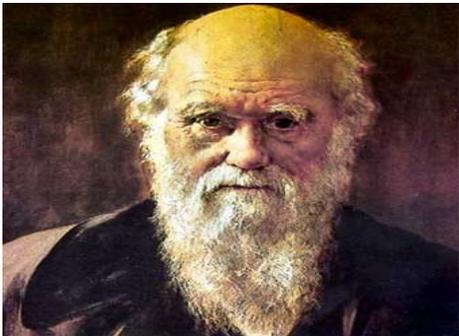


$$\min_{PI} \{ CAD(PI) \}$$

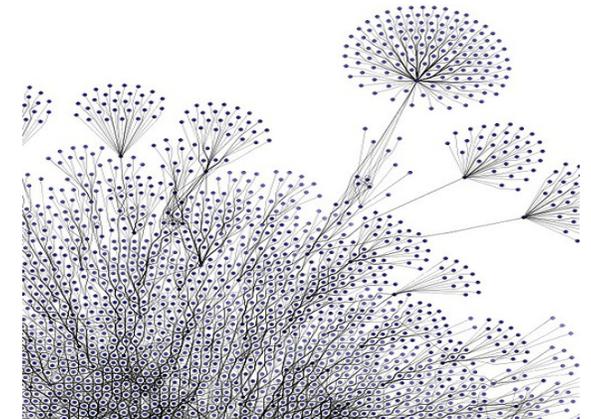
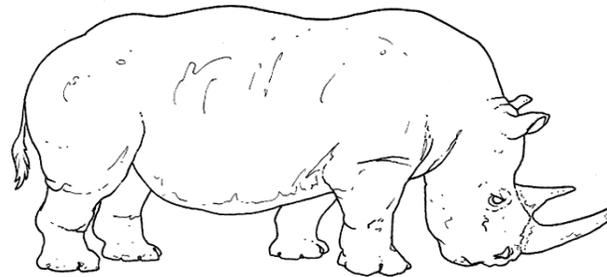
$$PI \in D$$

## Algoritmos genéticos

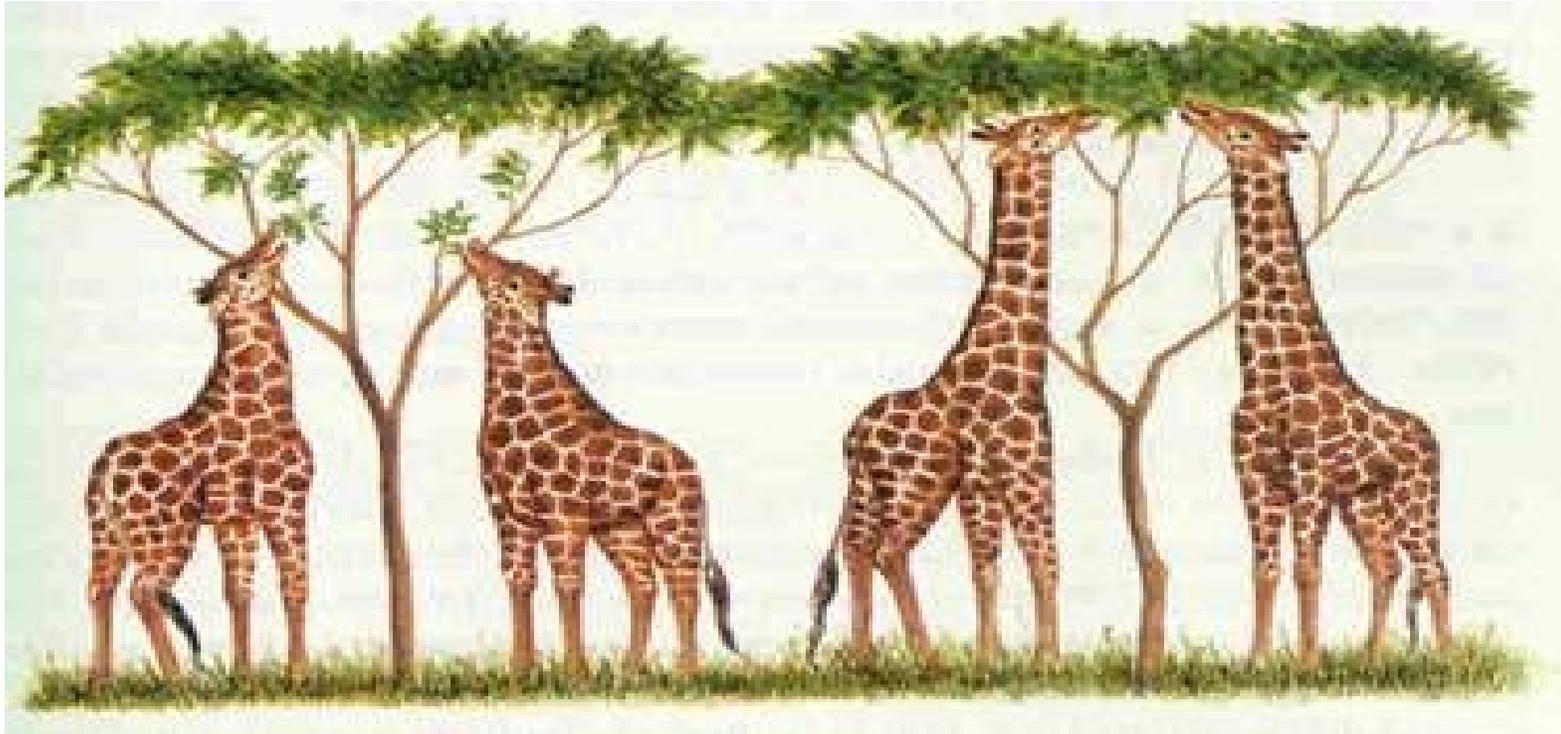
muchas variables + incertidumbres.



Principio de Selección Natural (1859)

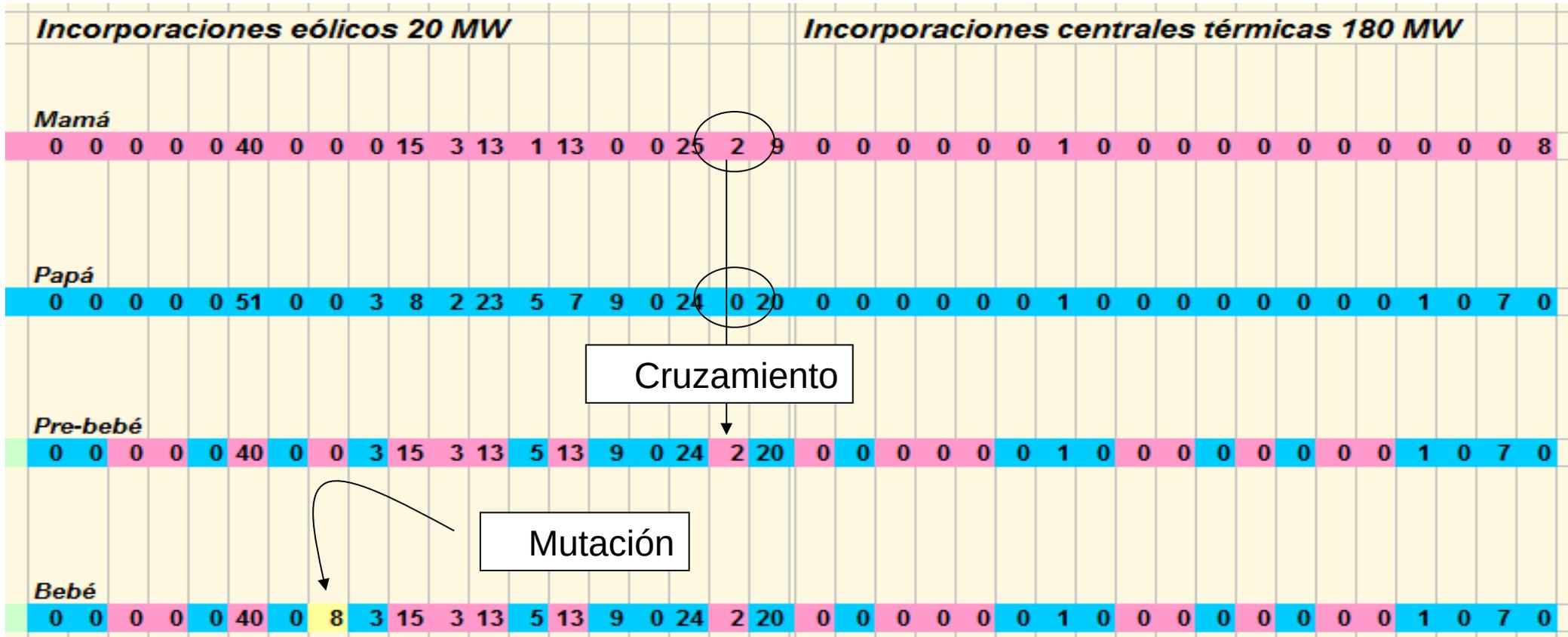


# Los más aptos...

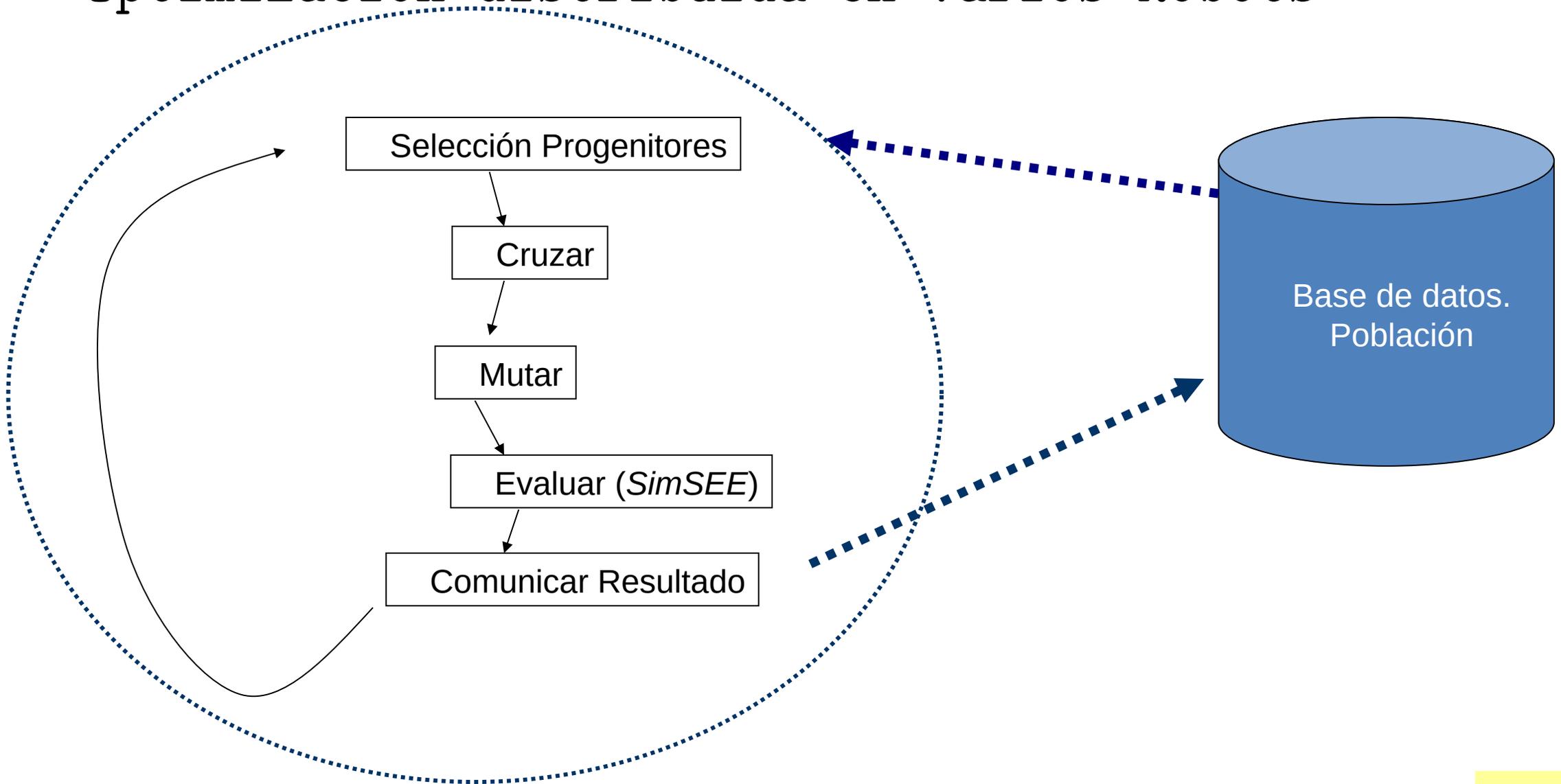


# Algoritmos genéticos

## Genotipo, cruzamiento y mutación



# Optimización distribuida en varios Robots



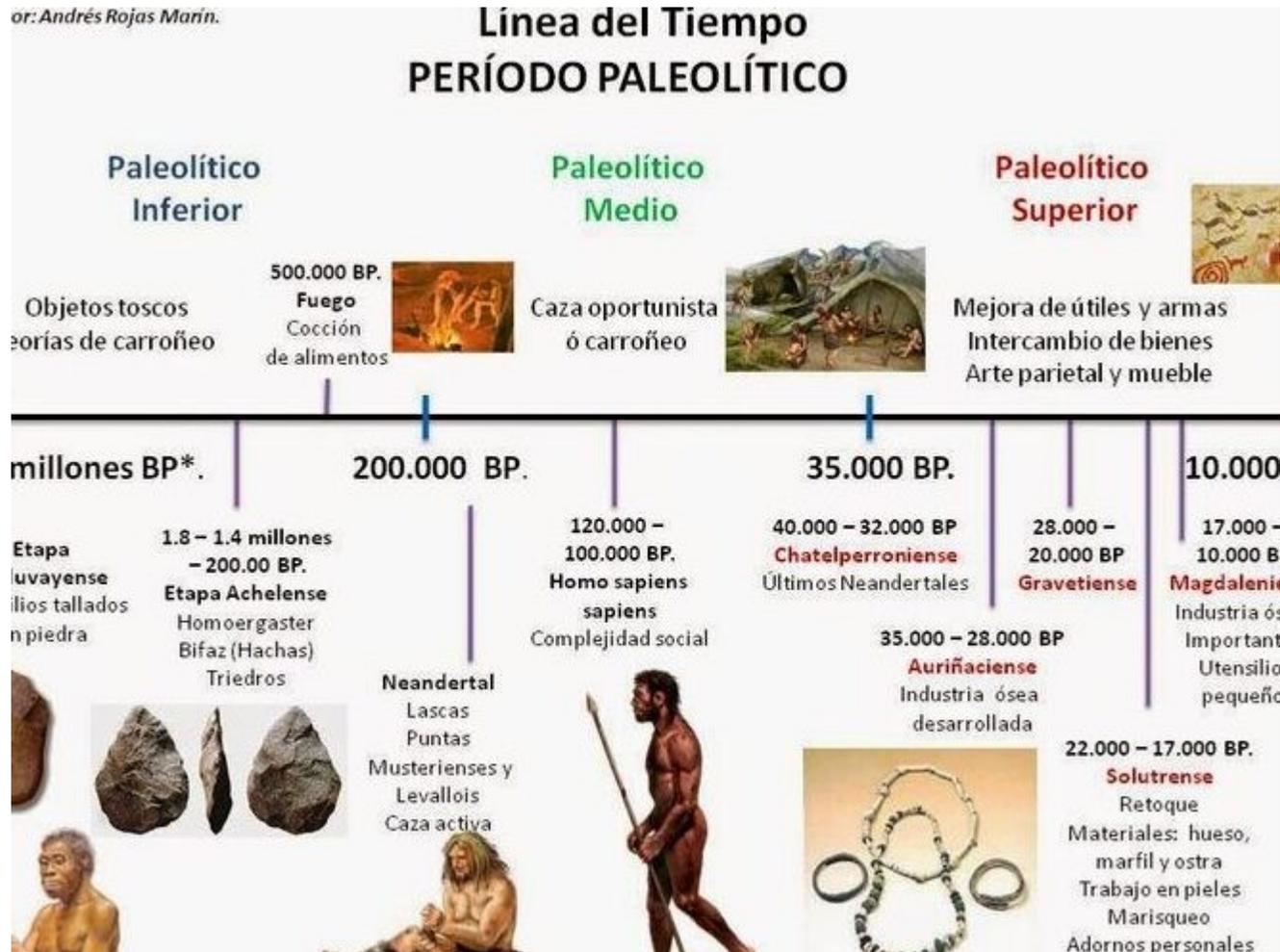
# Simulaciones Montecarlo en grupo de crónicas y re-evaluaciones



# ADN y Aprendizaje de la PO



or: Andrés Rojas Marín.



Con el mismo concepto de las crónicas de evaluación, se implementó la "oportunidad de aprendizaje".

El mismo ADN, ha tenido tantas oportunidades de evaluación como simulaciones y asociado al ADN tiene entonces una  $CF(X,k)$  correspondiente a su PO.

Cuando se cruzan individuos y se crea uno nuevo (que no ha sido evaluado aún) "nace" con el  $CF(X,k)$  de su Progenitor con menos evaluaciones. De esta forma, se va transmitiendo "lo aprendido" de generación en generación.



FIN

Muchas gracias por vuestra atención!