Extreme events on French Overseas territories

Physical mechanisms and development of forecasting tools at synoptic and subseasonal scales

Erwan Cornillault
Philippe Peyrillé, Romain Roehrig, Fleur Couvreux (CNRM - Météo-France)

17 Février 2026 - Journées IA de Météo-France

Extreme Precipitation Events

Extreme Precipitation Events (EPEs)

Rare and intense events
Major forecast issues and for **safety of people and property**
Numerous impacted socio-economic sectors

Between 2000 and 2019 globally (CRED & UNDRR, 2020) :

- **1.65 billion of people impacted** by floods, ~ 100.000 fatalities
- $**651 billions of damages**

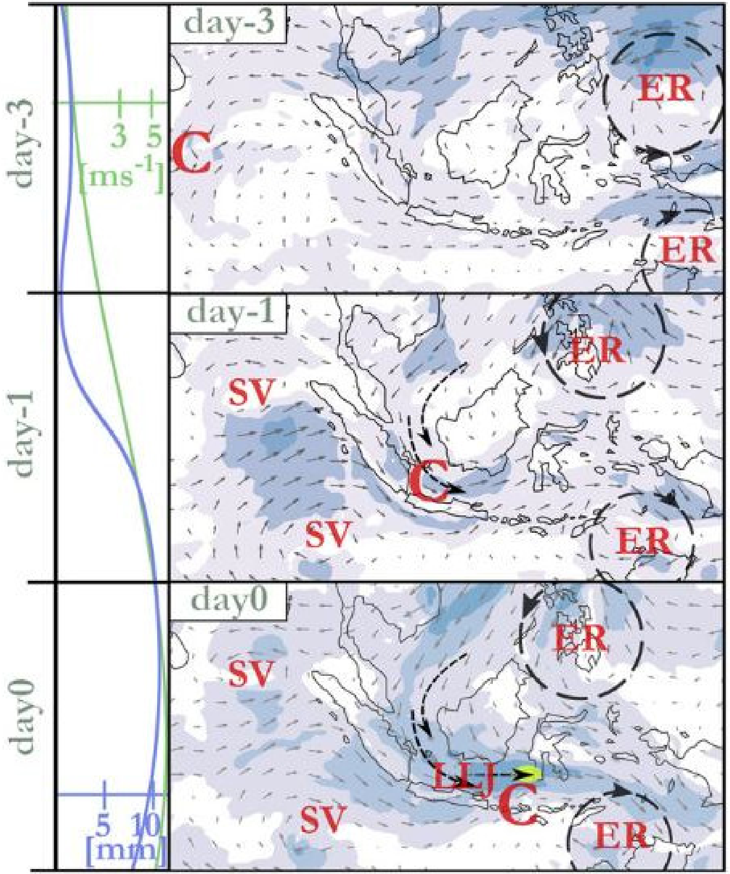
Issues

Poor skill of NPW models to forecast tropical EPEs

Tropical Cyclones : an important factor in EPEs but not the main one

Damages caused by EPEs in different OMs

EPEs and tropical variability



Rainfall, 850-hPa wind, and actors responsible for an EPE in Indonesia. C : low-level convergence of a Kelvin wave, ER : Equatorial Rossby waves, LLJ : low-level jet (Latos et al., *2021)*

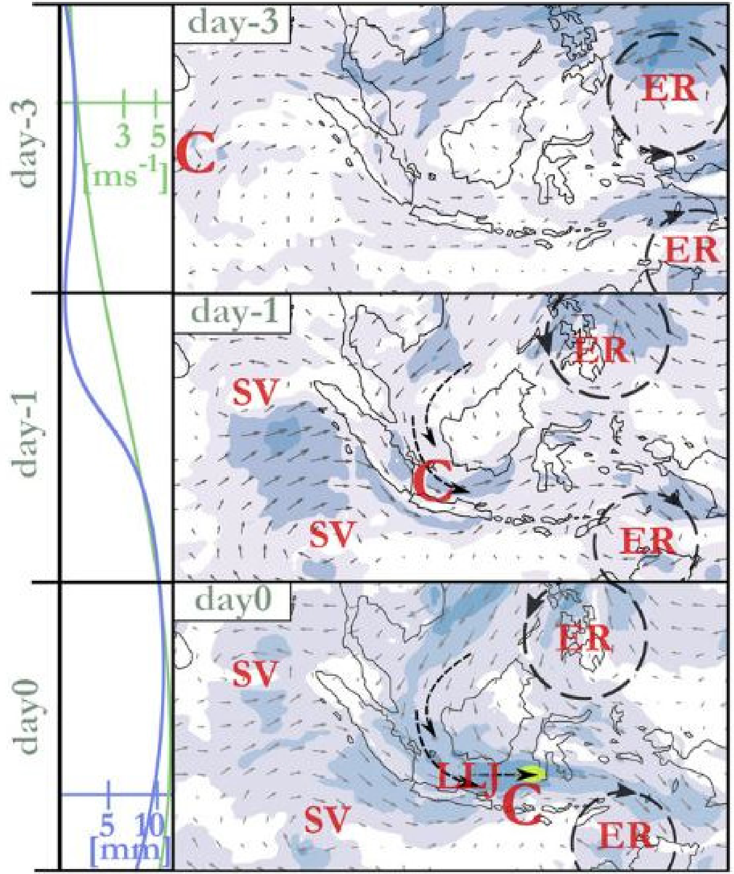
Equatorial waves

A few studies about the link between EPEs and intraseasonal variability (2-90 days)

EPE-prone situations : **superpositions of waves** (Latos et al., 2021; Peyrillé et al., 2023)

Mainly case studies, but **a few global and/or statistical studies about the link EPE / intraseasonal variability**

EPEs and tropical variability



Rainfall, 850-hPa wind, and actors responsible for an EPE in Indonesia. C : low-level convergence of a Kelvin wave, ER : Equatorial Rossby waves, LLJ : low-level jet (Latos et al., 2021)

Equatorial waves

A few studies about the link between EPEs and intraseasonal variability (2-90 days)

EPE-prone situations : **superpositions of waves** (Latos et al., 2021; Peyrillé et al., 2023)

Mainly case studies, but **a few global and/or statistical studies about the link EPE / intraseasonal variability**

What are the large-scale configurations associated with EPEs ?

Can we increase the EPE predictability thanks to large-scale indicators ?

Weather patterns in the tropics

Weather Patterns (WP)

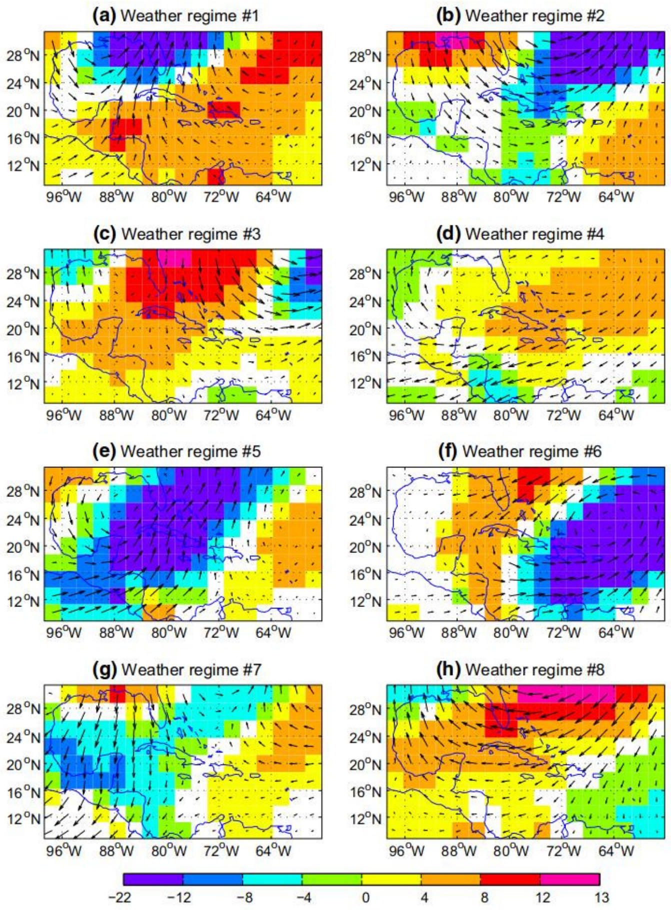
Large-scale and recurring atmospheric configurations, linked to local occurrences of rainfall, high or low temperatures, . . .

Similar large-scale configurations = same local impacts (Lorenz, 1969)

Several algorithms to identify WP (Grotjahn et al., 2016) : k*-means, PCA, hierarchical classifications, self-organizing maps, . . .*

Weather patterns in the tropics

OLR (shadings in W/m²) and 925-hPa winds for eight WPs across the Caribbean Sea, expressed as anomalies. (Moron et al., 2016)



Weather Patterns (WP)

Large-scale and recurring atmospheric configurations, linked to local occurrences of rainfall, high or low temperatures, ...

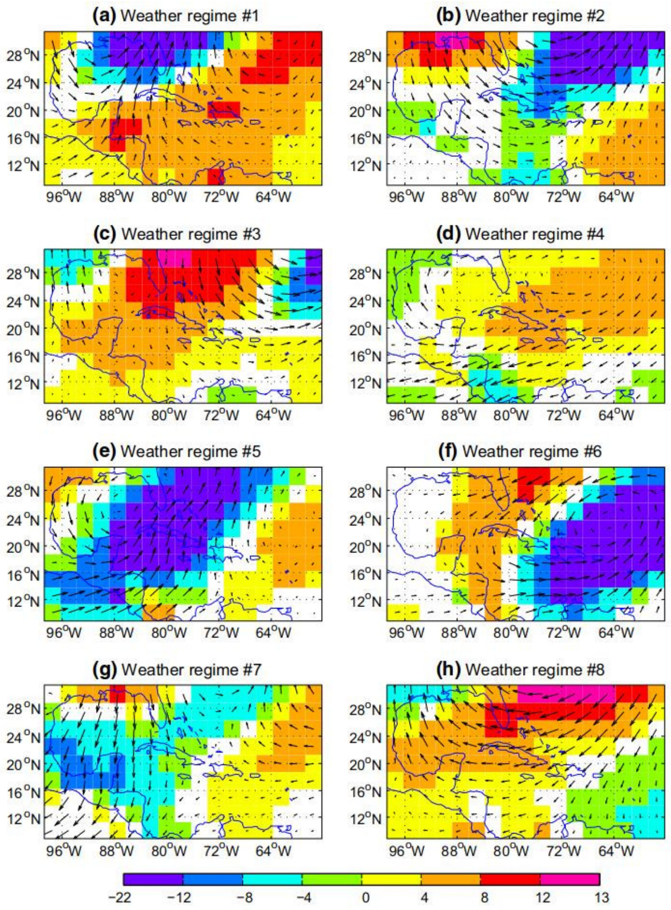
Similar large-scale configurations = same local impacts (Lorenz, 1969)

Several algorithms to identify WP (Grotjahn et al., 2016) : *k-means, PCA, hierarchical classifications, self-organizing maps, ...*

Few studies in the tropics, but growing literature (e.g., Moron et al., 2015, 2016 ; Rapolaki et al., 2019 ; Réchou et al., 2019)

Weather patterns in the tropics

OLR (shadings in W/m^2) and 925-hPa winds for eight WPs across the Caribbean Sea, expressed as anomalies. (Moron et al., 2016)



Weather Patterns (WP)

Large-scale and recurring atmospheric configurations, linked to local occurrences of rainfall, high or low temperatures, ...

Similar large-scale configurations = same local impacts (Lorenz, 1969)

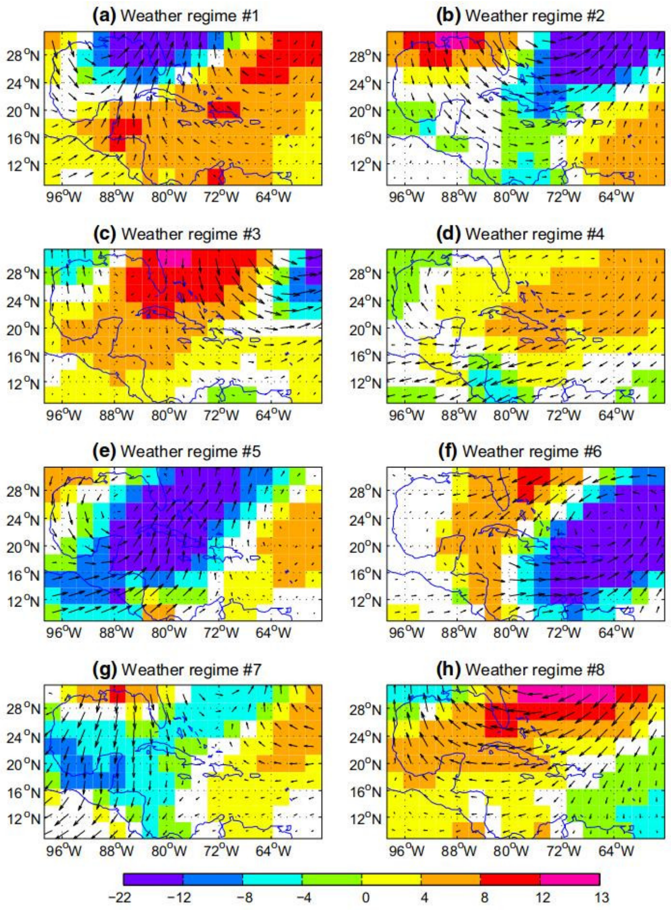
Several algorithms to identify WP (Grotjahn et al., 2016) : *k-means, PCA, hierarchical classifications, self-organizing maps, ...*

Few studies in the tropics, but growing literature (e.g., Moron et al., 2015, 2016 ; Rapolaki et al., 2019 ; Réchou et al., 2019)

Can EPE-related weather patterns be defined over French Overseas areas ?

Weather patterns in the tropics

OLR (shadings in W/m^2) and 925-hPa winds for eight WPs across the Caribbean Sea, expressed as anomalies. (Moron et al., 2016)



Weather Patterns (WP)

Large-scale and recurring atmospheric configurations, linked to local occurrences of rainfall, high or low temperatures, . . .

Similar large-scale configurations = same local impacts (Lorenz, 1969)

Several algorithms to identify WP (Grotjahn et al., 2016) : *k-means, PCA, hierarchical classifications, self-organizing maps, . . .*

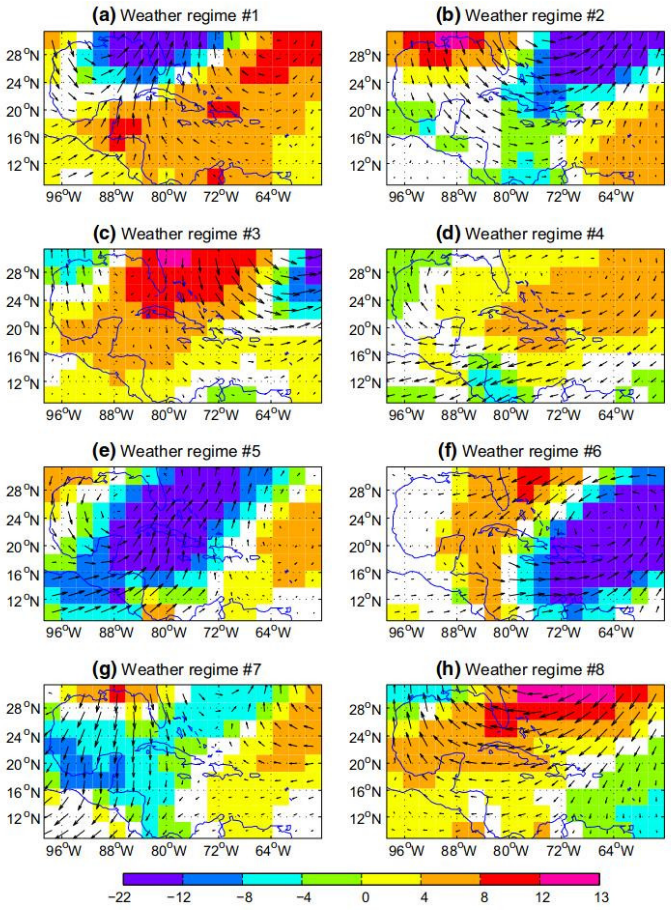
Few studies in the tropics, but growing literature (e.g., Moron et al., 2015, 2016 ; Rapolaki et al., 2019 ; Réchou et al., 2019)

Can EPE-related weather patterns be defined over French Overseas areas ?

Nguyen-Le et Yamada (2019) et Nguyen-Le et al. (2017) : analog forecast of EPEs in Japan and Thailand, thanks to WP.

Weather patterns in the tropics

OLR (shadings in W/m²) and 925-hPa winds for eight WPs across the Caribbean Sea, expressed as anomalies. (Moron et al., 2016)



Weather Patterns (WP)

Large-scale and recurring atmospheric configurations, linked to local occurrences of rainfall, high or low temperatures, . . .

Similar large-scale configurations = same local impacts (Lorenz, 1969)

Several algorithms to identify WP (Grotjahn et al., 2016) : k-*means, PCA, hierarchical classifications, self-organizing maps, . . .*

Few studies in the tropics, but growing literature (e.g., Moron et al., 2015, 2016 ; Rapolaki et al., 2019 ; Réchou et al., 2019)

Can EPE-related weather patterns be defined over French Overseas areas ?

Nguyen-Le et Yamada (2019) et Nguyen-Le et al. (2017) : analog forecast of EPEs in Japan and Thailand, thanks to WP.

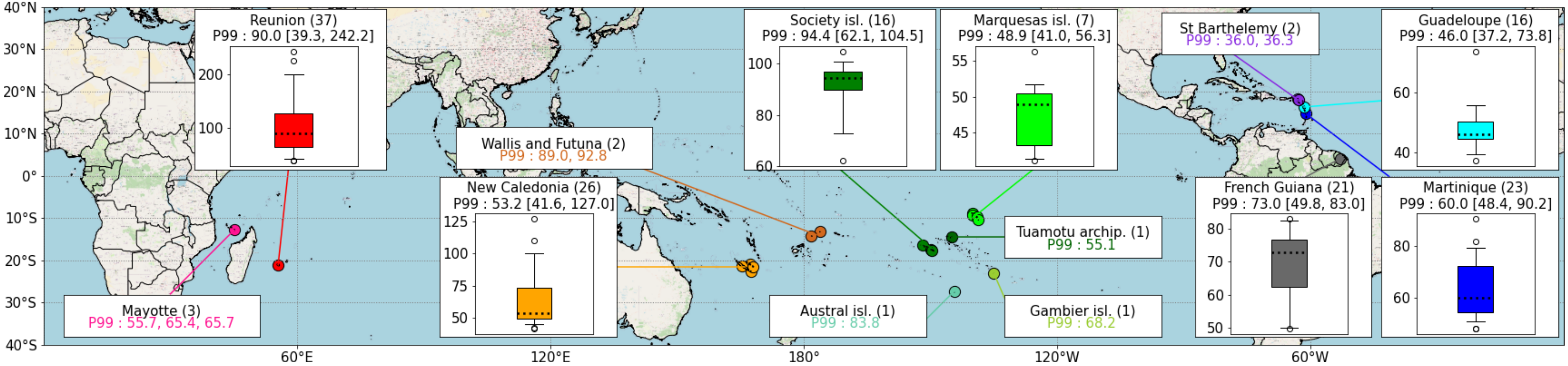
Can EPE predictability be increased in OMs, thanks to WPs ?

Data and EPE definition

Rainfall data
- **Daily rainfall** on the period **1979-2021**
- Rain gauges with less than 5 % of missing data

→ 156 rain gauges, some of which are above 1000 meters

EPE definition
- **RR24 ≥ p99 (all-day)** on each station
- **150 to 160 EPEs** for each station. **Cyclonic EPEs** are withdrawn.

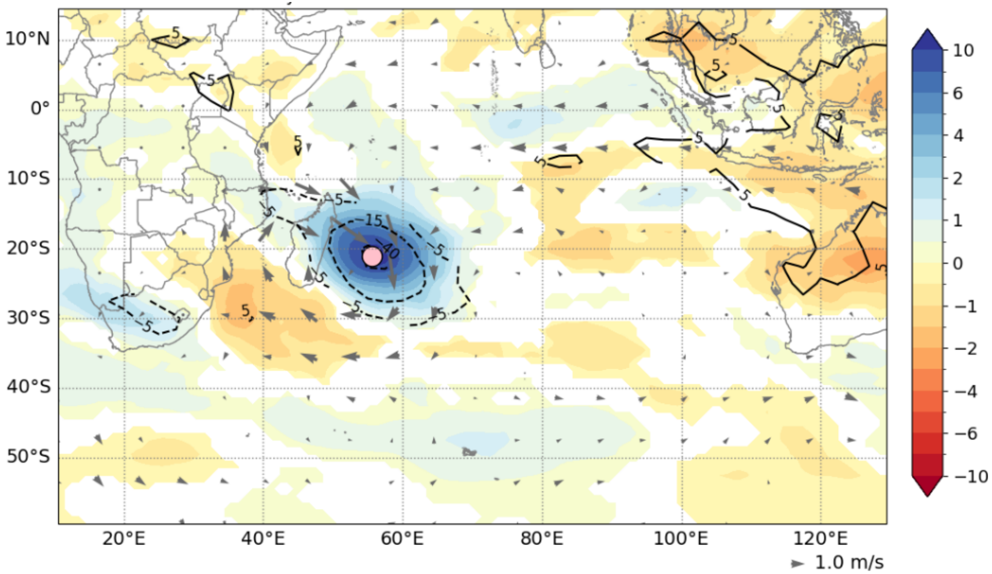


Different thresholds depending on the rain gauge to **take into account the high variability of rainfall climatology** (*inter-OM* and *intra-OM* variability)

Large-scale mean environment of EPEs

What are the main configurations responsible for EPEs in OMs ?
What scale impacts them the most ?

Large-scale mean environment of EPEs



Composites anomalies of PW (colors), OLR (contours), and 850-hPa wind (arrows) during EPEs in Réunion

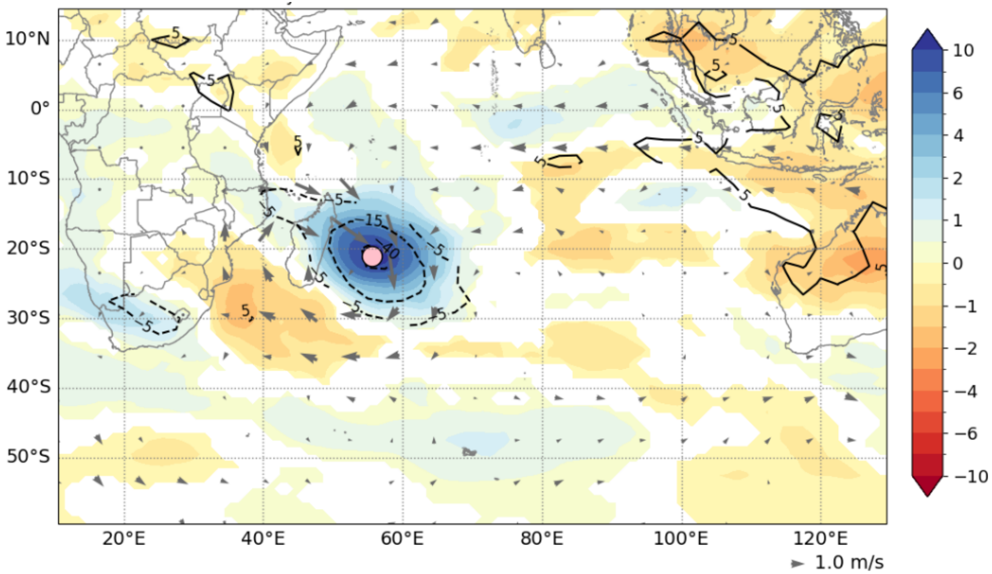
Données

- PW, 850-hPa wind (ERA5 1°, Hersbach et al., 2020)
- OLR (NOAA, 2.5°, Liebmann & Smith, 1996)

In Réunion :

- **Moist and convective anomaly**, a few thousand of kilometres
- Cyclonic anomaly close to Réunion

Large-scale mean environment of EPEs



Composites anomalies of PW (colors), OLR (contours), and 850-hPa wind (arrows) during EPEs in Réunion

Données

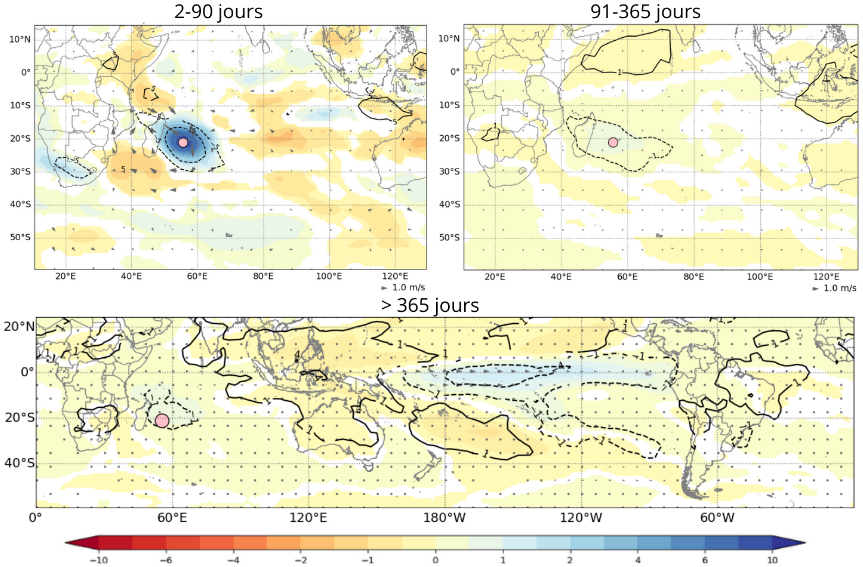
- PW, 850-hPa wind (ERA5 1°, Hersbach et al., 2020)
- OLR (NOAA, 2.5°, Liebmann & Smith, 1996)

In Réunion :

- **Moist and convective anomaly**, a few thousand of kilometres
- Cyclonic anomaly close to Réunion

Close behaviour between all OMs

Contribution of typical temporal scales

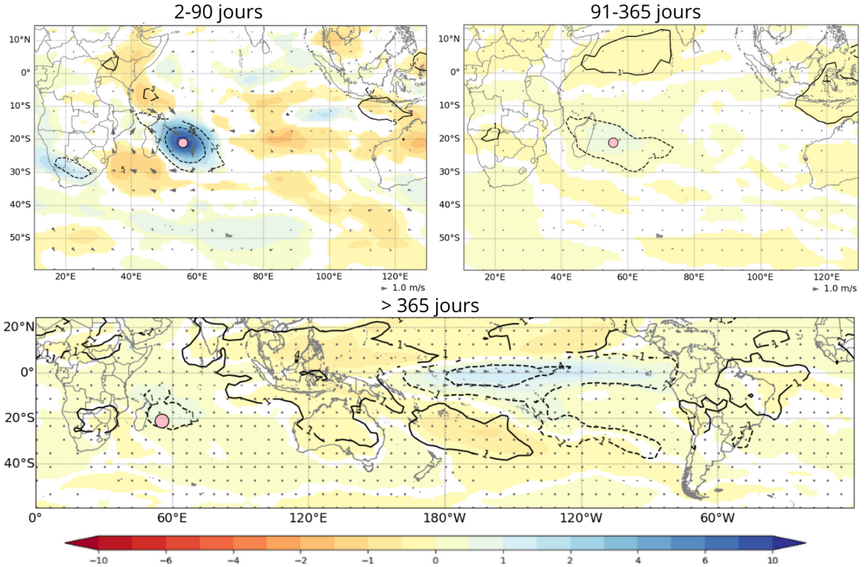


Intraseasonal, seasonal-to-annual and interannual composite anomalies during d'EPEs in Réunion

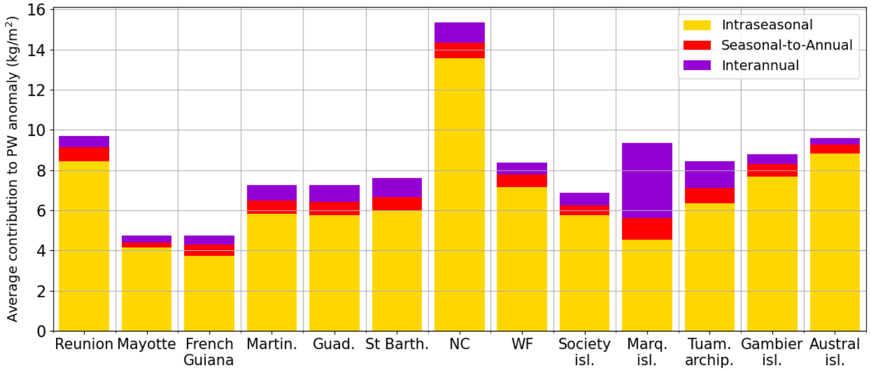
In Réunion :

- **Intraseasonal = main contribution**
- Weak contribution of other scales (one order of magnitude lower)

Contribution of typical temporal scales



Intraseasonal, seasonal-to-annual and interannual composite anomalies during d'EPEs in Réunion



Contribution of the three scales to the PW total anomaly during EPEs

Pour tous les OMs :

- Contribution majoritaire de l'intrasaisonnier : **fortes anomalies intrasaisonnières ⇒ fortes anomalies totales**
- Signal venant en grande partie des ondes équatoriales

In Réunion :

- **Intraseasonal = main contribution**
- Weak contribution of other scales (one order of magnitude lower)

Cornillault, E., P. Peyrille, F. Couvreux, and R. Roehrig. 2024. "Large-Scale Drivers of Tropical Extreme Precipitation Events : The Example of French Overseas Territories". Geophysical Research Letters 51 (15) : e2024GL108770. https://doi.org/10.1029/2024GL108770.

Classification of EPE large-scale environment

Can we identify typical large-scale situations associated with EPEs ?

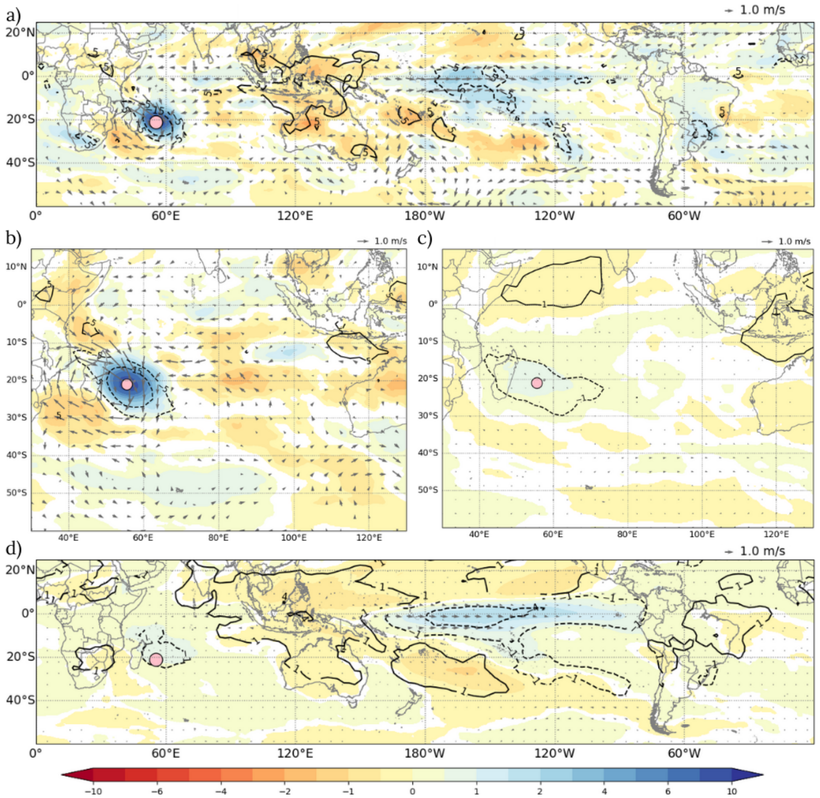
What are these situations ?

Atmospheric training variables

Represent humidity and circulation that can impact EPEs$_{nc}$ occurrence

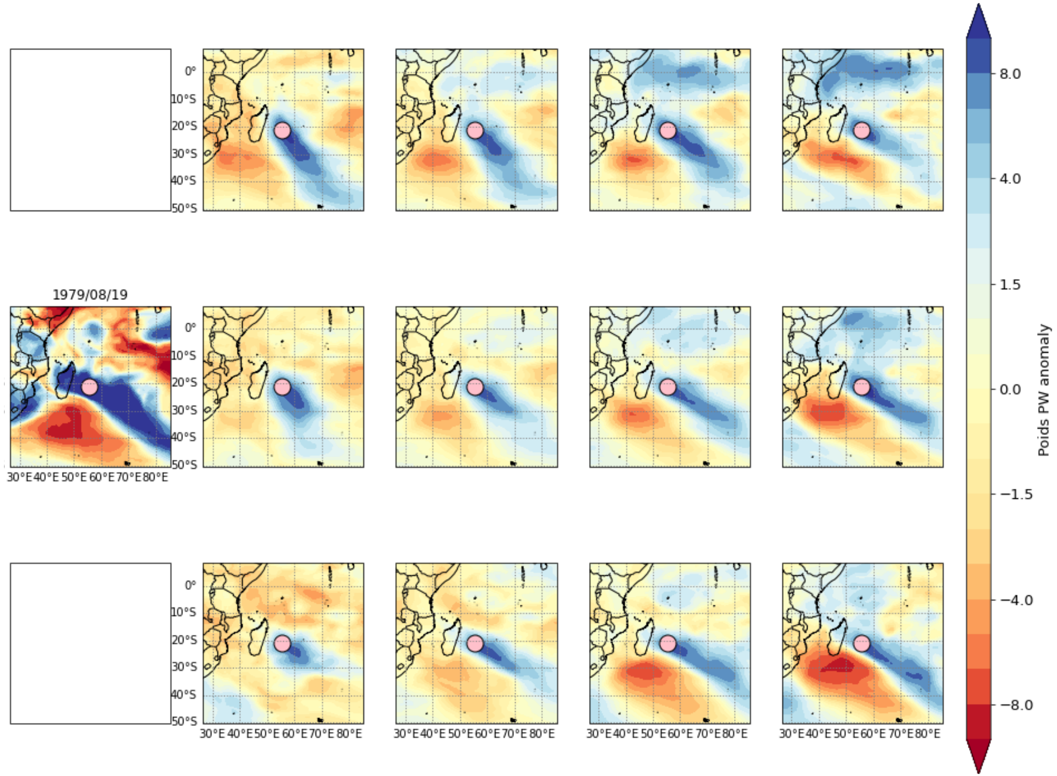
- **200 and 850 hPa wind** : description of low-level (trade winds, monsoon, ...) and high-level circulation (jets, ageostrophic forcings)
- **PW** : a proxy for tropical rainfall, and on average, correlates with EPEs and a very humid environment around territories during EPEs$_{nc}$ (Cornillault et al., 2024)
- **Raw and anomaly fields** : seasonality and difference with climatology
- **CCEW-filtered anomalies** : major variability modes, intraseasonal anomalies contribute the most to total anomalies, source of predictability

⇒ 30 variables with (almost) no linear combination between them



Composite anomalies of PW, OLR, and 850 hPa wind during EPEs$_{nc}$ in La Réunion (Cornillault et al., 2024). a : total anomalies. b,c,d : 2-90 day, 91-365 day, and 365+ day filtered anomalies

Self-Organizing Maps



Looking for patterns that synthesize all possible situations that lead to an EPE_{nc} : clusters can have heterogeneous sizes.

- PCA initialisation
- Euclidean distance between nodes and data

Illustration of SOM training. PW anomaly and situation of 1979/08/19.

Self-Organizing Maps

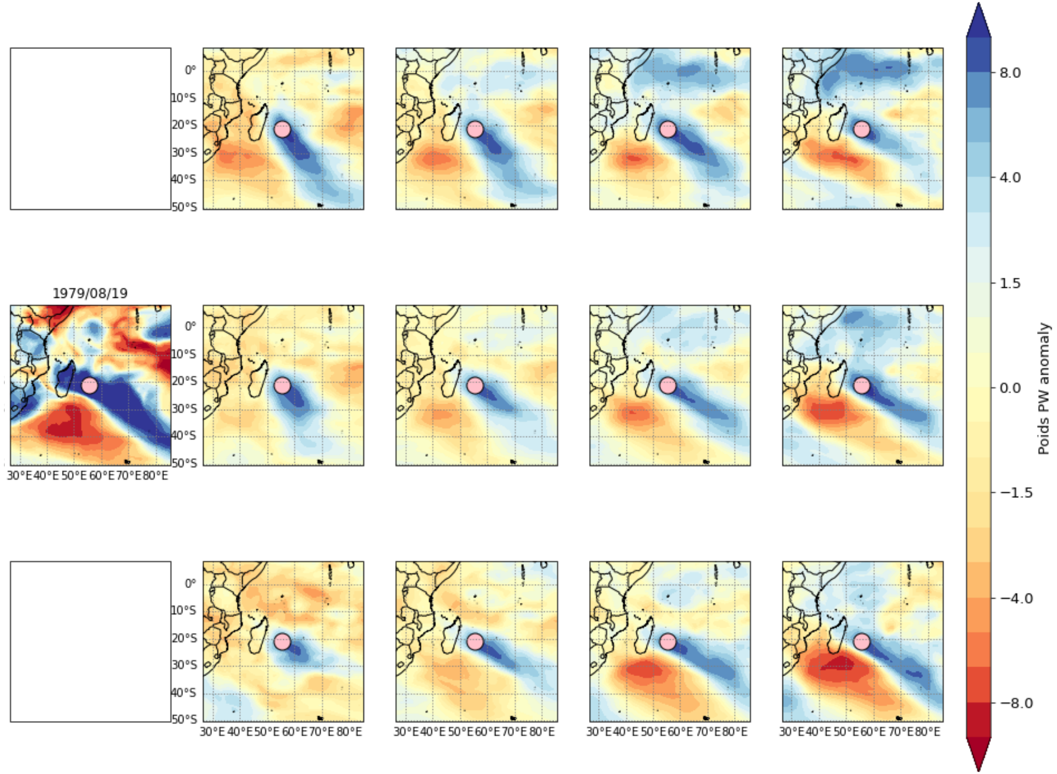


Illustration of SOM training. PW anomaly and situation of 1979/08/19.

Looking for patterns that synthesize all possible situations that lead to an EPE_{nc} : clusters can have heterogeneous sizes.

- PCA initialisation
- Euclidean distance between nodes and data
- Best Matching Unit : closest node with data (smaller distance)

Self-Organizing Maps

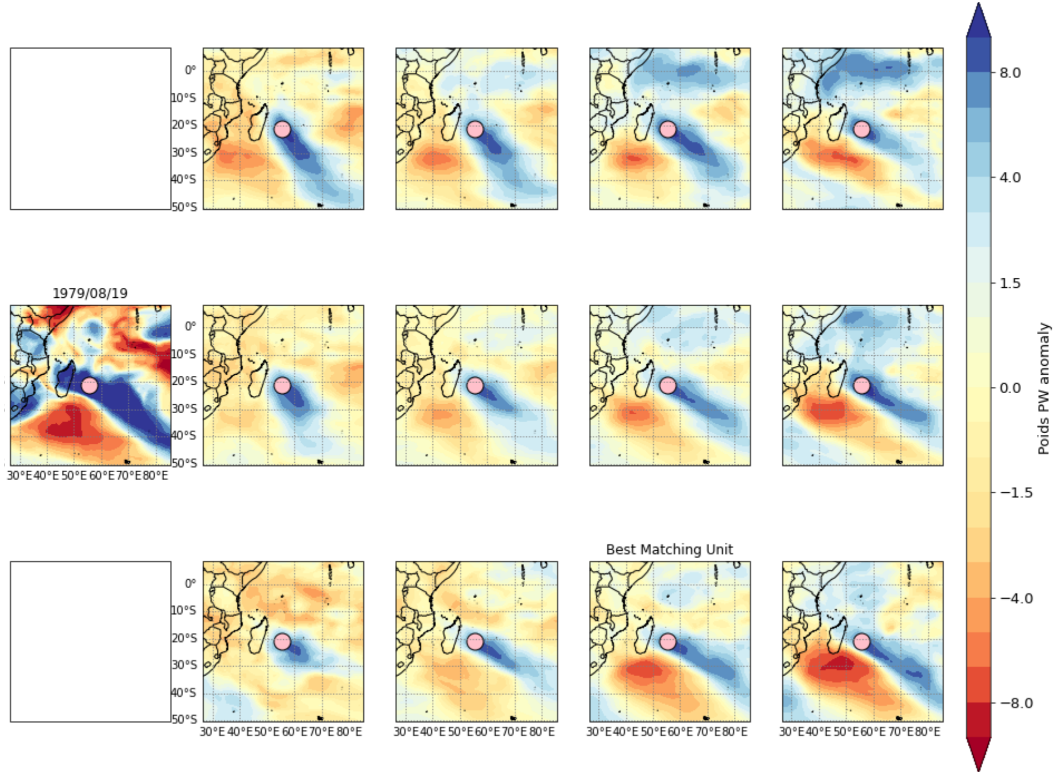


Illustration of SOM training. PW anomaly and situation of 1979/08/19.

Looking for patterns that synthesize all possible situations that lead to an EPE$_{nc}$: clusters can have heterogeneous sizes.

- PCA initialisation
- Euclidean distance between nodes and data
- Best Matching Unit : closest node with data (smaller distance)
- Update with a learning function, weighted by a neighbouring function. Gradual reduction of the interaction between neighbouring neurons to converge.

Self-Organizing Maps

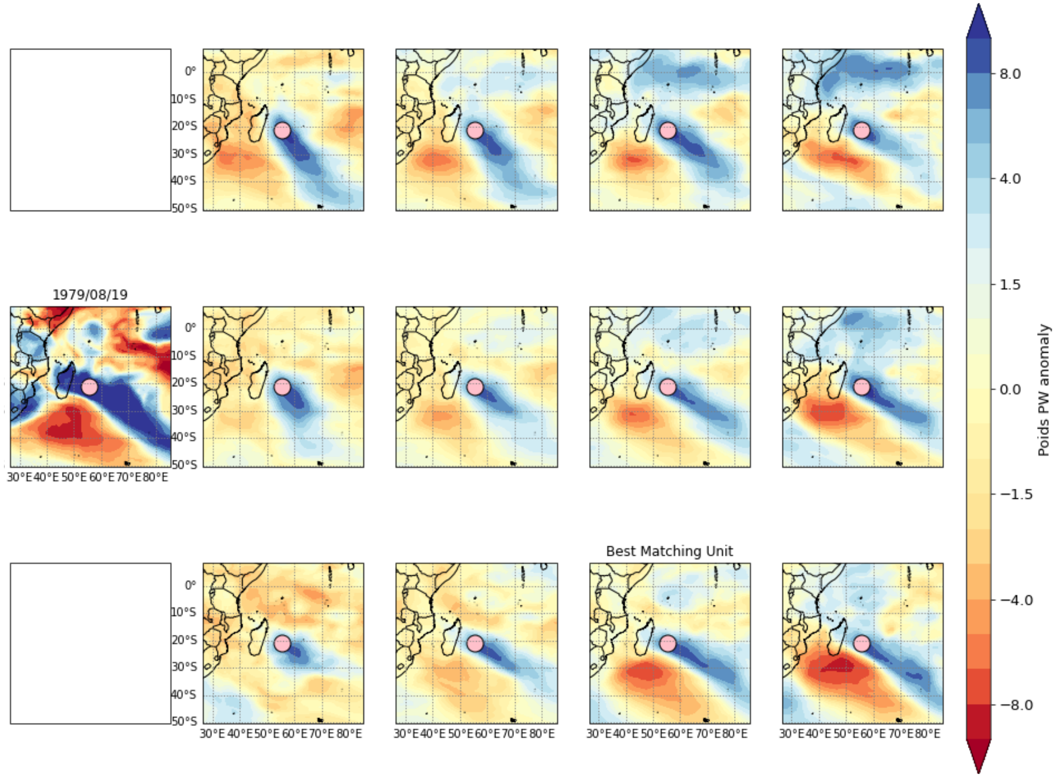


Illustration of SOM training. PW anomaly and situation of 1979/08/19.

Looking for patterns that synthesize all possible situations that lead to an EPE$_{nc}$: clusters can have heterogeneous sizes.

- PCA initialisation
- Euclidean distance between nodes and data
- Best Matching Unit : closest node with data (smaller distance)
- Update with a learning function, weighted by a neighbouring function. Gradual reduction of the interaction between neighbouring neurons to converge.

⇒ **Order** in the network, continuous distribution of WPs
⇒ **Preserving the topological structure** of the data
⇒ **Dimension reduction**

Self-Organizing Maps

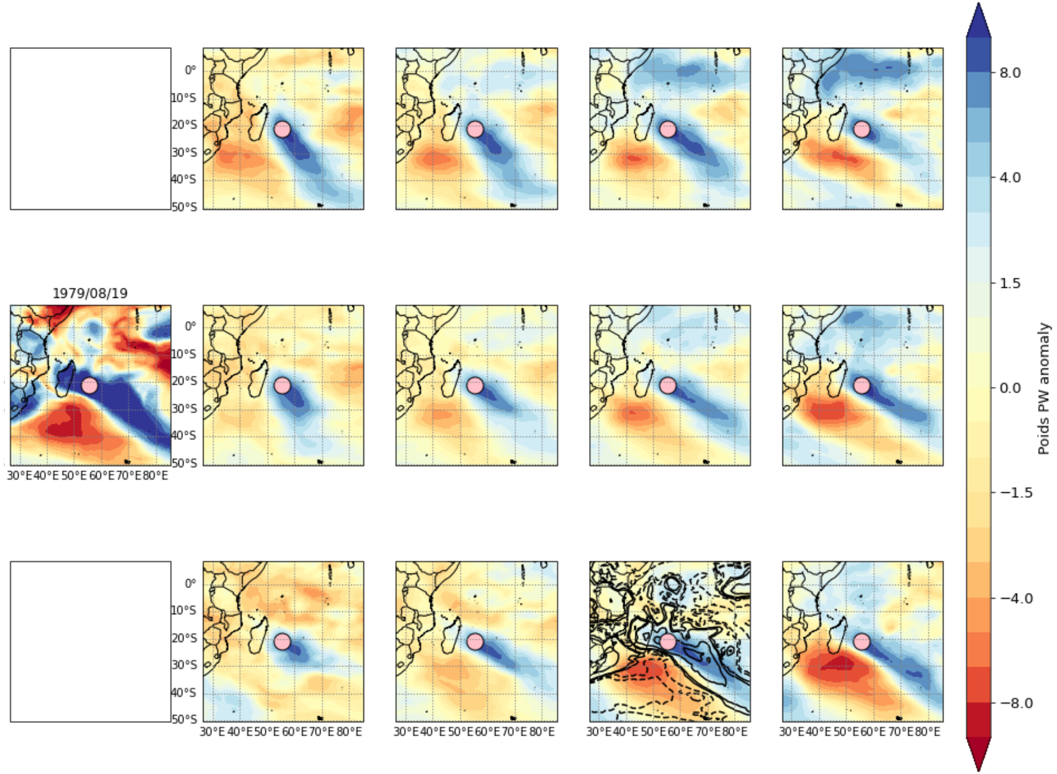


Illustration of SOM training. PW anomaly and situation of 1979/08/19.

Looking for patterns that synthesize all possible situations that lead to an EPE$_{nc}$: clusters can have heterogeneous sizes.

- PCA initialisation
- Euclidean distance between nodes and data
- Best Matching Unit : closest node with data (smaller distance)
- Update with a learning function, weighted by a neighbouring function. Gradual reduction of the interaction between neighbouring neurons to converge.

⇒ **Order** in the network, continuous distribution of WPs
⇒ **Preserving the topological structure** of the data
⇒ **Dimension reduction**

Self-Organizing Maps

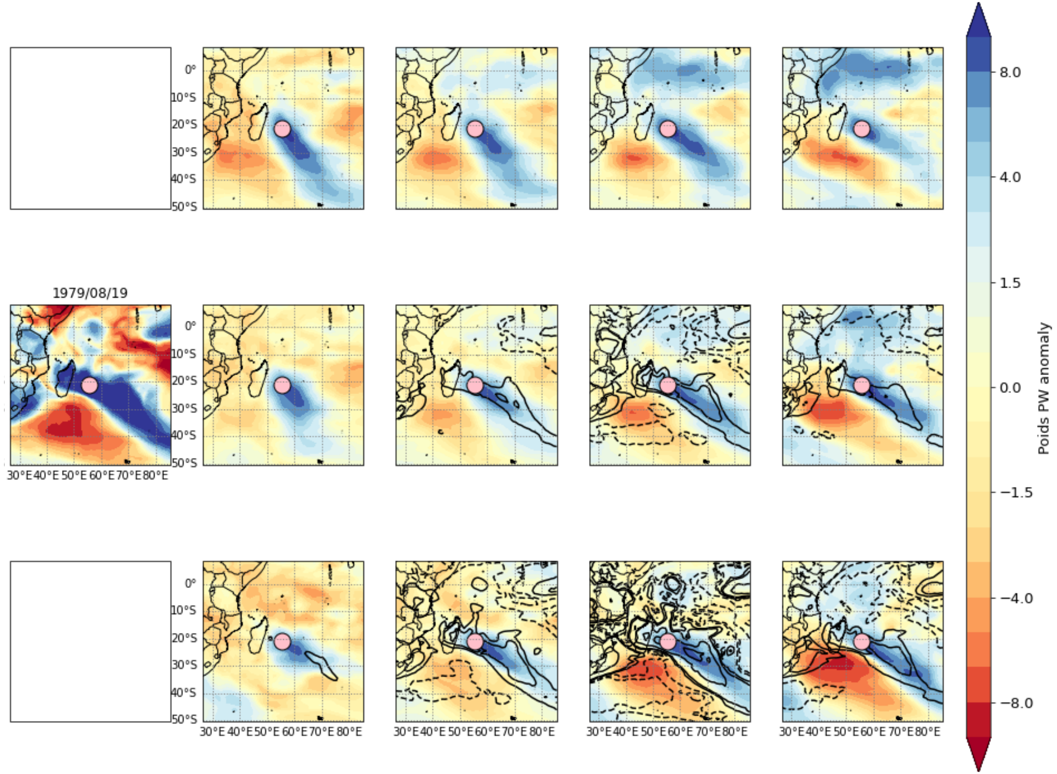


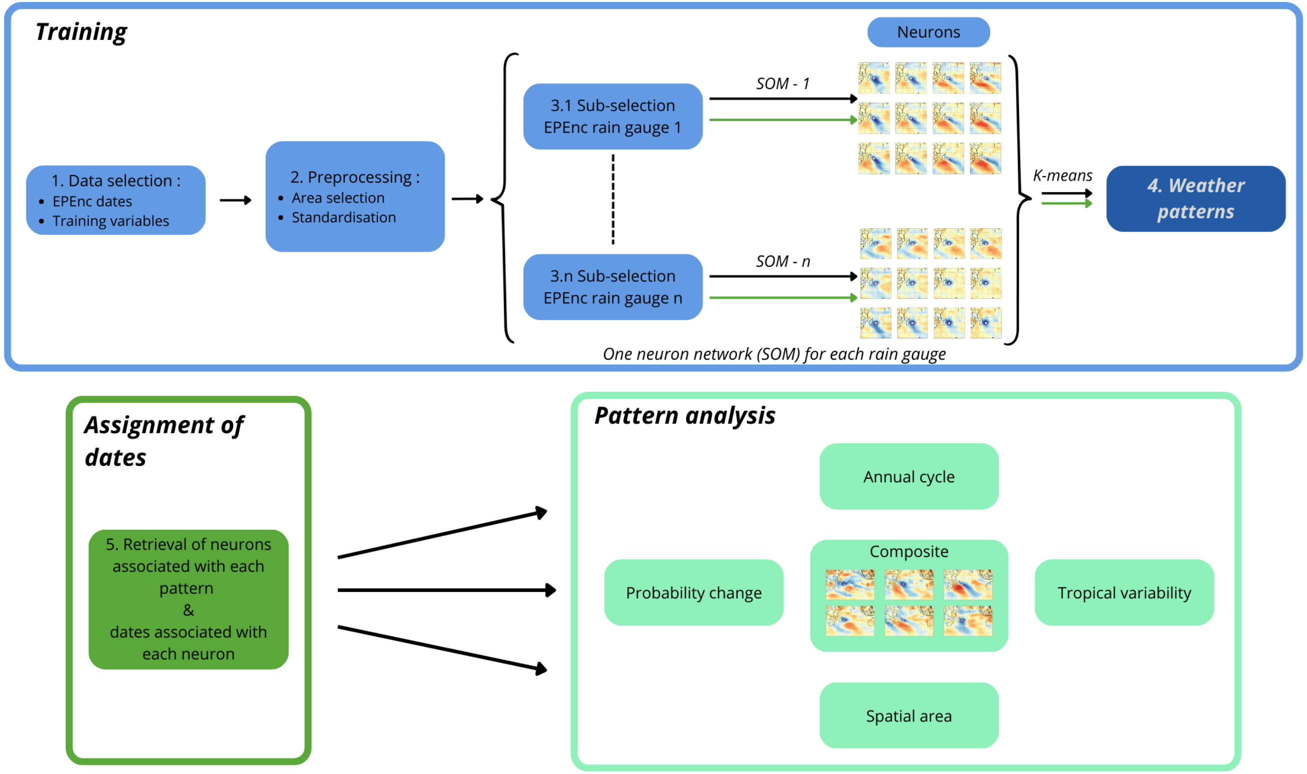
Illustration of SOM training. PW anomaly and situation of 1979/08/19.

Looking for patterns that synthesize all possible situations that lead to an EPE_{nc} : clusters can have heterogeneous sizes.

- PCA initialisation
- Euclidean distance between nodes and data
- Best Matching Unit : closest node with data (smaller distance)
- Update with a learning function, weighted by a neighbouring function. Gradual reduction of the interaction between neighbouring neurons to converge.

⇒ **Order** in the network, continuous distribution of WPs
⇒ **Preserving the topological structure** of the data
⇒ **Dimension reduction**

Method of WP classification



Numerous sensitivity tests to adjust hyperparameters and the training method

Common method for all territories (only difference : final number of WPs) and currently fully completed only for La Réunion.

Sensitivity tests : SOMs

Evaluation metrics (average over all SOMs) :

- **Quantization error** : quadratic error between nodes and data
- **Topographic error** : fraction of data whose the 1st and 2nd BMUs are not neighbours (network order)
- **Number of empty nodes** : average number of nodes which is not a BMU for any data

Training mode and number of iterations

- `online` : one random vector is used to update the map per iteration
- `batch` : all data are used to update the map per iteration

Sensitivity tests : SOMs

Evaluation metrics (average over all SOMs) :

- **Quantization error** : quadratic error between nodes and data
- **Topographic error** : fraction of data whose the 1st and 2nd BMUs are not neighbours (network order)
- **Number of empty nodes** : average number of nodes which is not a BMU for any data

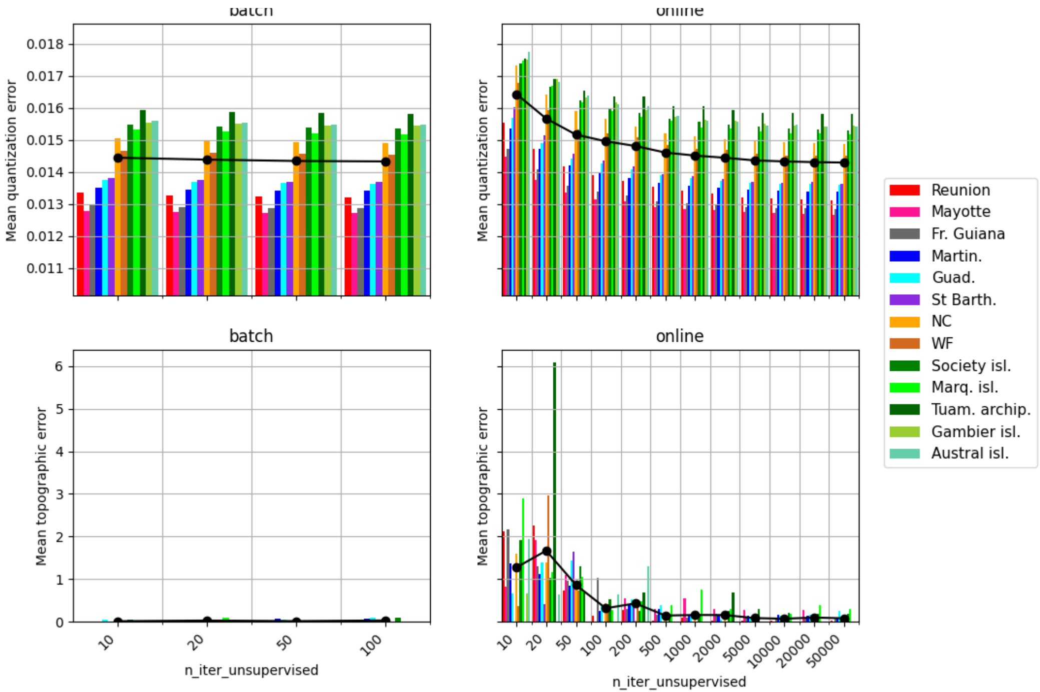
Training mode and number of iterations

- `online` : one random vector is used to update the map per iteration
- `batch` : all data are used to update the map per iteration

For the same number of iterations :

- higher memory cost and longer training time (10 à 70× more) with a `batch` training
- faster convergence with a `batch` training → 10 à 100× faster to reach the same errors
- null topographic error with a `batch` training

⇒ **`batch` training with 50 iterations**



SOM error sensitivity to the training mode as a function of the number of iterations

Sensitivity tests : SOMs

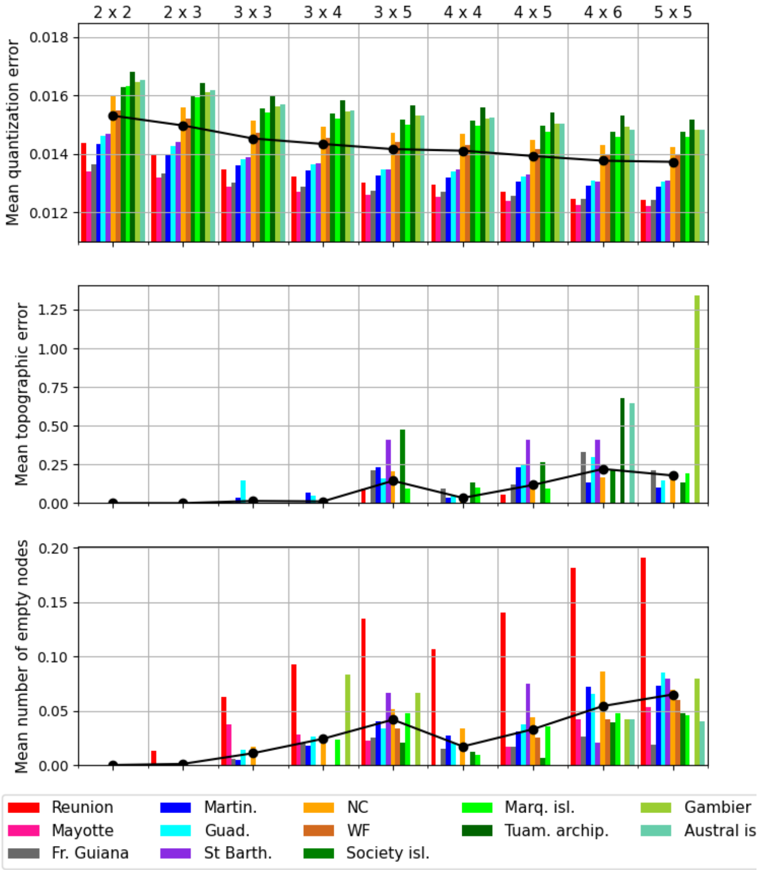
Evaluation metrics (average over all SOMs) :

- **Quantization error** : quadratic error between nodes and data
- **Topographic error** : fraction of data whose the 1st and 2nd BMUs are not neighbours (network order)

Number of nodes and shape of the map :

- lower quantization error with more nodes
- higher topographic error and number of empty nodes with more nodes

⇒ compromise between the errors. Choice : **3×4 map**



SOM error sensitivity to the number of nodes and the shape for each territory

Sensitivity tests : k-means

Evaluation metrics :

- **Silhouette** : score based on the vicinity between each piece of data and each cluster, to quantify how much data is well classified
- **Proportion of correct classification** : fraction of data with a positive silhouette.
- **Davies-Bouldin index** : measure of the separation between clusters and the vicinity inside clusters. Lower = better classification

Sensitivity tests : k-means

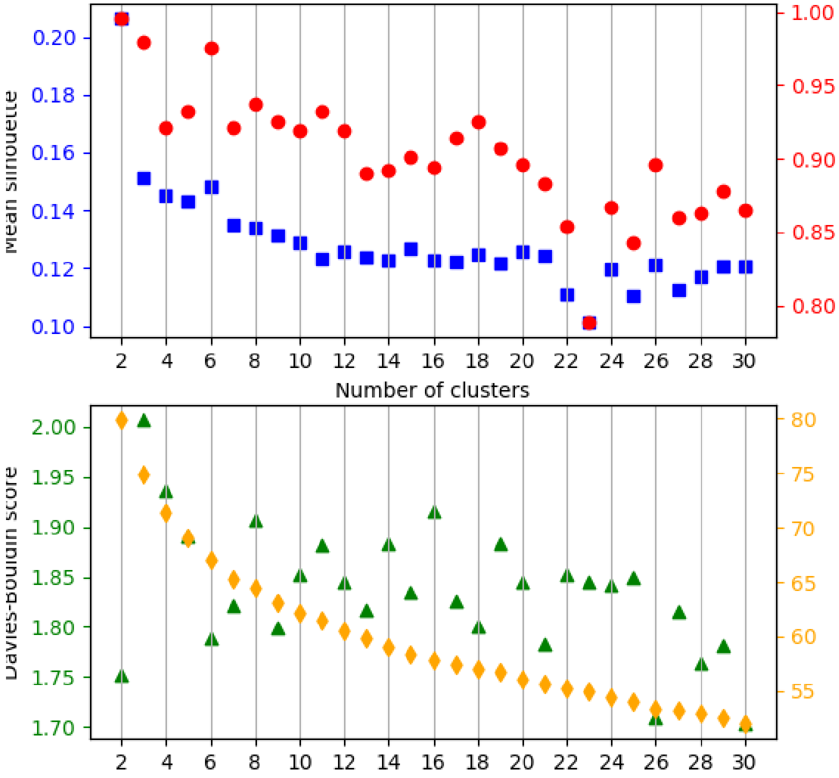
Evaluation metrics :

- **Silhouette** : score based on the vicinity between each piece of data and each cluster, to quantify how much data is well classified
- **Proportion of correct classification** : fraction of data with a positive silhouette.
- **Davies-Bouldin index** : measure of the separation between clusters and the vicinity inside clusters. Lower = better classification

Number of WPs that maximize silhouette / PCC and minimize DB index and RMSE :

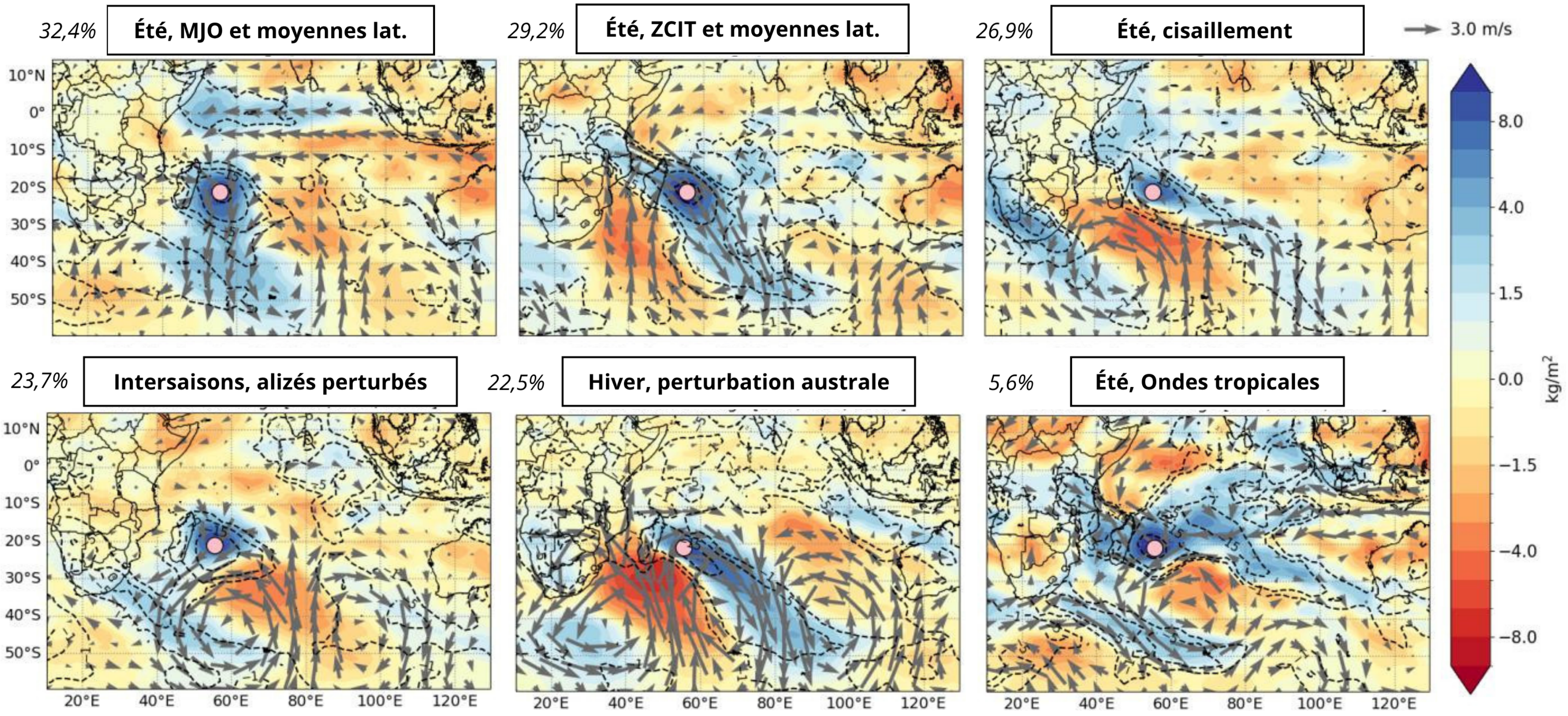
- several possible choices : 6, 15, 18, 26
- Need to be easily interpretable by a human forecaster (no black box) = a low number of WPs is better.

⇒ **6 weather patterns for EPEs$_{nc}$ in La Réunion**



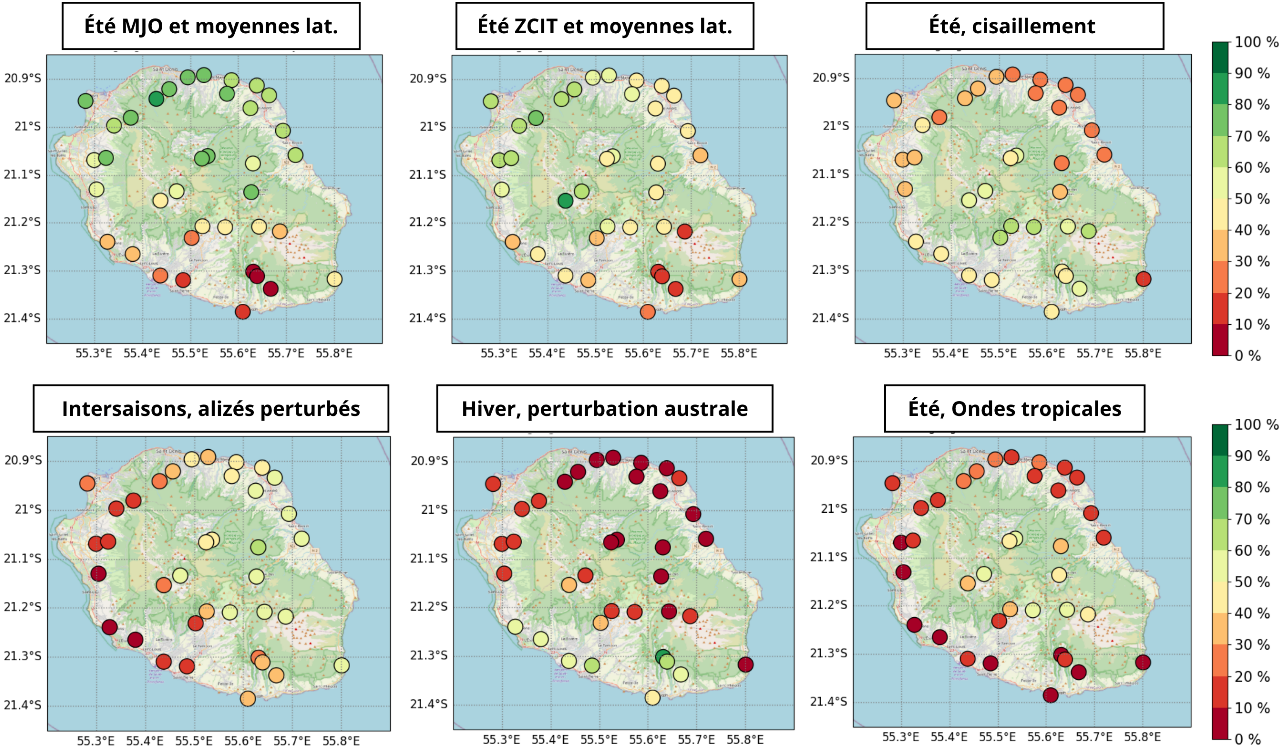
Classification scores as a function of the final number of WPs, for La Réunion. Top : silhouette (blue) and PCC (red). Bottom : Davies-Bouldin index (green) and RMSE (yellow).

EPE Weather Patterns in Réunion



PW / OLR/ 850-hPa composite anomalies for each WP and fraction of occurrence

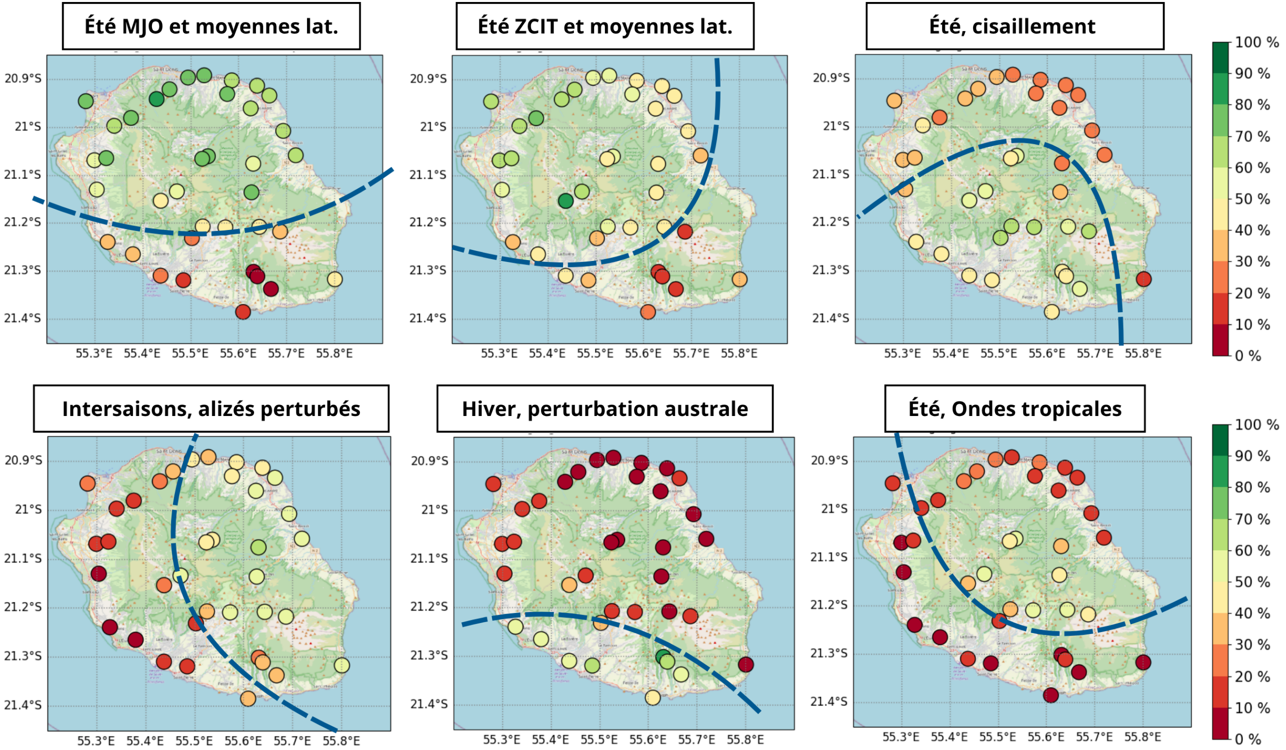
Fraction of EPEs associated with each WP



$\forall g$ a rain gauge

$$\text{Fraction of EPEs (g)} = \frac{nb_{EPE} \,/\, WP(g)}{nb_{EPE}\ total(g)}$$

Fraction of EPEs associated with each WP

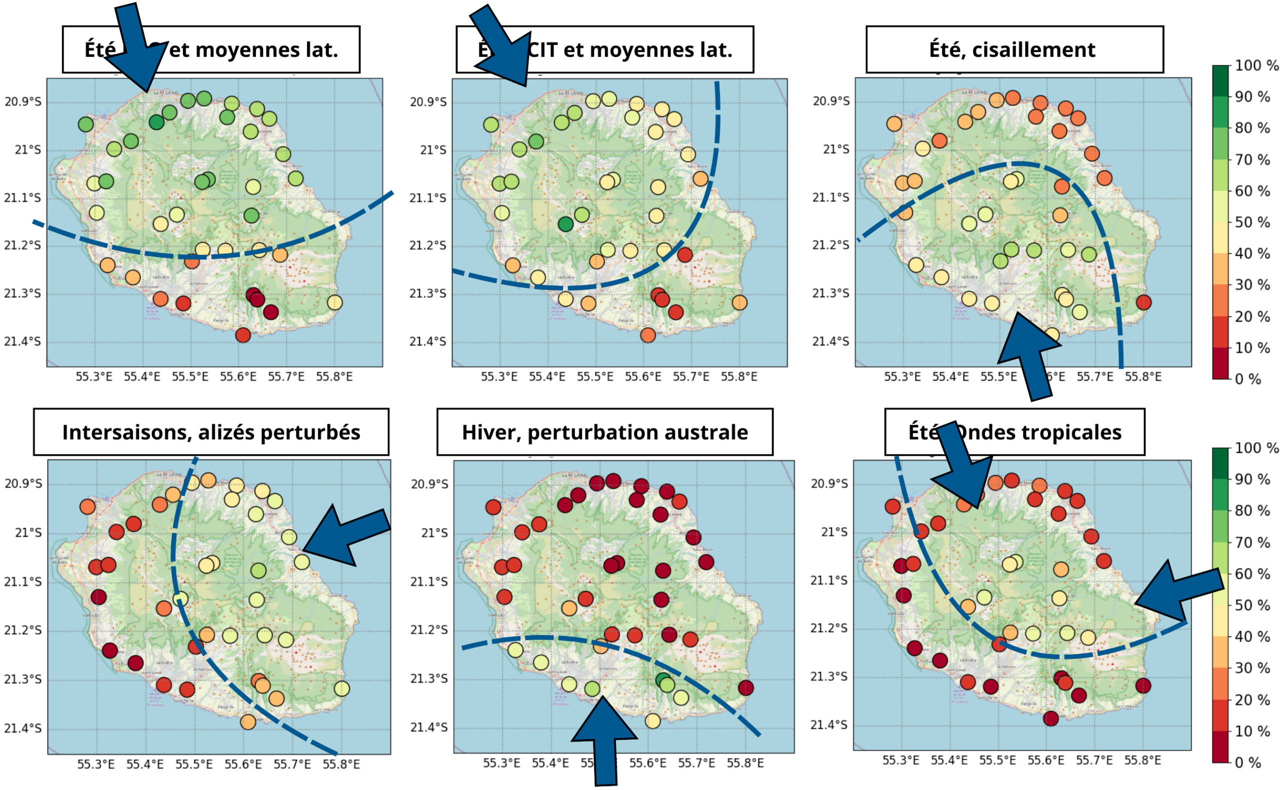


$\forall g$ a rain gauge

$$\text{Fraction of EPEs (g)} = \frac{\text{nb}_{EPE} \ / \ \text{WP}(g)}{\text{nb}_{EPE} \ \text{total}(g)}$$

- **Specific area of EPE occurrence** for each WP

Fraction of EPEs associated with each WP



$\forall g$ a rain gauge

$$\text{Fraction of EPEs (g)} = \frac{\text{nb}_{EPE} \;/\; \text{WP}(g)}{\text{nb}_{EPE}\;\text{total}(g)}$$

- **Specific area of EPE occurrence** for each WP
- In accordance with the **direction of low-level wind** → orographic lift

Impact of large-scale weather patterns at the local scale

EPE predictability

Can EPE predictability be increased thanks to weather patterns ?

Analog forecast (Lorenz, 1969)

What criterion for analog ?

Multivariate linear correlation (all training variables) on the geographical training area between two situations or one situation Q and one WP R

$$c_{Q,R} = \frac{Cov(Q,R)}{\sigma_Q.\sigma_R}$$

Two analog situations = high $c_{Q,R}$

From what threshold of $c_{Q,R}$ a situation Q is analog to R ?

Analog forecast (Lorenz, 1969)

What criterion for analog ?

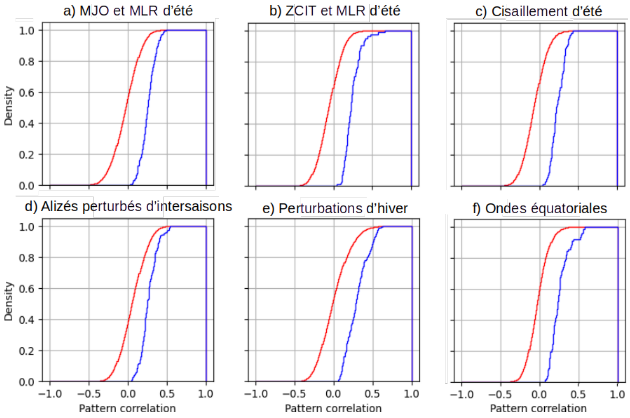
Multivariate linear correlation (all training variables) on the geographical training area between two situations or one situation Q and one WP R

$$c_{Q,R} = \frac{Cov(Q,R)}{\sigma_Q.\sigma_R}$$

Two analog situations = high $c_{Q,R}$

From what threshold of $c_{Q,R}$ a situation Q is analog to R ?

CDF of $c_{Q,R}$ between each WP and dates between 1979 and 2021 (all dates and EPEs)



Analog forecast (Lorenz, 1969)

What criterion for analog ?

Multivariate linear correlation (all training variables) on the geographical training area between two situations or one situation Q and one WP R

$$c_{Q,R} = \frac{Cov(Q,R)}{\sigma_Q.\sigma_R}$$

Two analog situations = high $c_{Q,R}$

From what threshold of $c_{Q,R}$ a situation Q is analog to R ?

CDF of $c_{Q,R}$ between each WP and dates between 1979 and 2021 (all dates and EPEs)

Contingency table

Count of **occurrences or not, observed and forecast EPE** :

- ***Hit rate***, H : $\frac{a}{a+c}$ (to maximize)
- ***False alarm rate*** (F) : $\frac{b}{b+d}$ (to minimize)

		Observed		
		Yes	No	
Forecast	Yes	a	b	a + b
	No	c	d	c + d
		a + c	b + d	n = a + b + c + d

Analog forecast (Lorenz, 1969)

What criterion for analog ?

Multivariate linear correlation (all training variables) on the geographical training area between two situations or one situation Q and one WP R

$$c_{Q,R} = \frac{Cov(Q,R)}{\sigma_Q.\sigma_R}$$

Two analog situations = high $c_{Q,R}$

From what threshold of $c_{Q,R}$ a situation Q is analog to R ?

CDF of $c_{Q,R}$ between each WP and dates between 1979 and 2021 (all dates and EPEs)

Contingency table

Count of **occurrences or not, observed and forecast EPE** :

- ***Hit rate***, H : $\frac{a}{a+c}$ (to maximize)
- ***False alarm rate*** (F) : $\frac{b}{b+d}$ (to minimize)

		Observed		
		Yes	No	
Forecast	Yes	a	b	a + b
	No	c	d	c + d
		a + c	b + d	n = a + b + c + d

Sensitivity of H and F with the correlation threshold

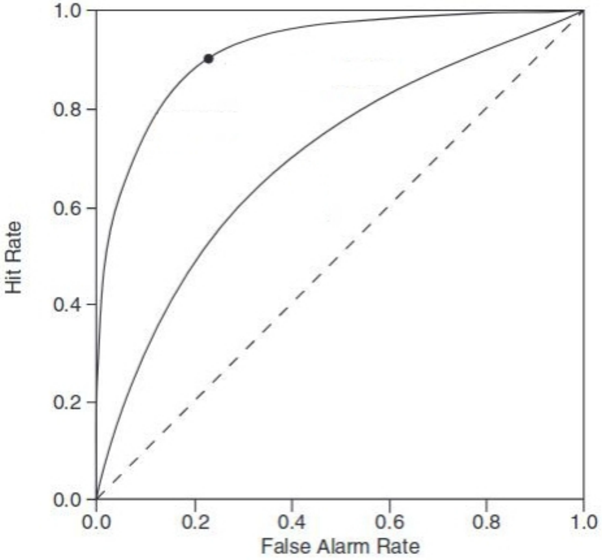
For each increment of $c \in [-1, 1]$, filling the contingency table

- if $c_{Q,R} > c \Rightarrow$ forecast of an occurrence, overwise non-occurrence
- Observed EPE = situation Q with at least one EPE for one rain gauge

Skill of analog forecast

ROC Diagram (*Receiver Operating Characteristic*)
Hit rate (H) and **False alarm rate** (F) depending on the threshold c

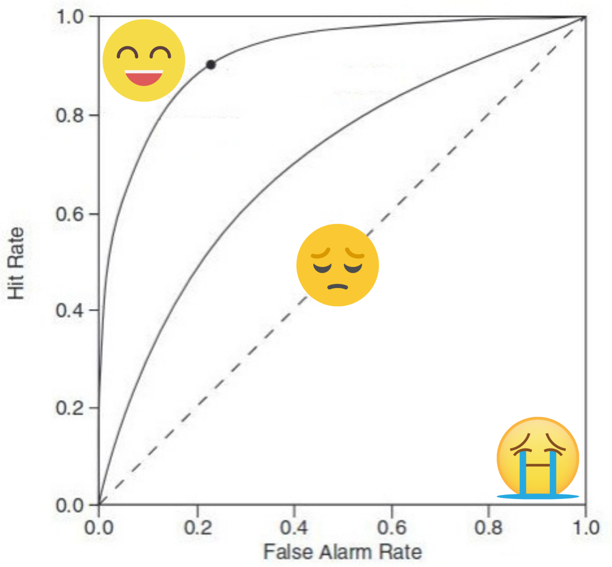
Good model = ROC curve away from the diagonal
Optimal $c_{Q,R}$ = the closest point to (1, 0)



Skill of analog forecast

ROC Diagram (*Receiver Operating Characteristic*)
Hit rate (H) and **False alarm rate** (F) depending on the threshold c

Good model = ROC curve away from the diagonal
Optimal $c_{Q,R}$ = the closest point to (1, 0)

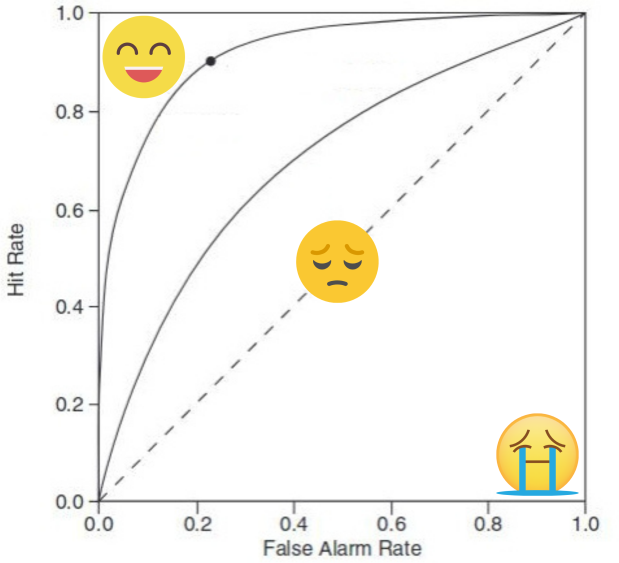


⇒ Evaluate the **predictability potential** on the climatology and find the **optimal threshold of** $c_{Q,R}$

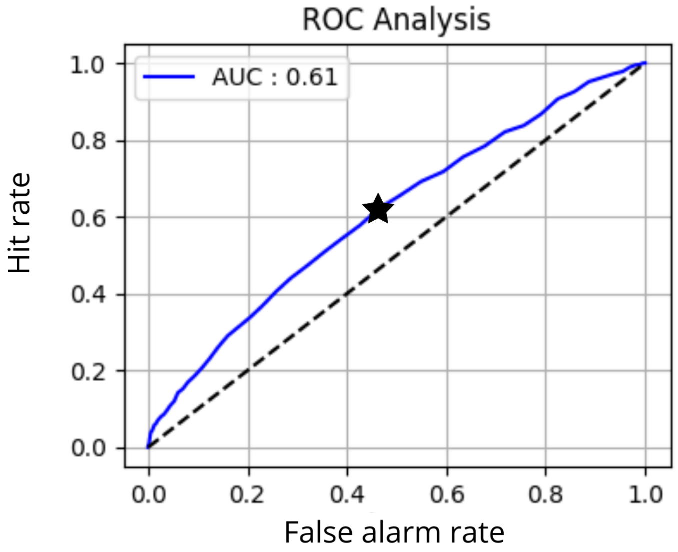
Skill of analog forecast

ROC Diagram (*Receiver Operating Characteristic*)
Hit rate (H) and **False alarm rate** (F) depending on the threshold c

Good model = ROC curve away from the diagonal
Optimal $c_{Q,R}$ = the closest point to (1, 0)



⇒ Evaluate the **predictability potential** on the climatology and find the **optimal threshold of** $c_{Q,R}$



ROC analysis for Réunion :

- ROC curve > diagonal
- Optimal $c_{Q,R}$: H = 0.51, F = 0.35 ⇒ 58 % of EPE occurrence probability if an EPE is forecast

Little predictability potential, better model than a random model

Skill of analog forecast with a temporal tolerance

n-day temporal tolerance

If an EPE is forecast between $J-n$ and $J+n$, is it observed between $J-n$ and $J+n$?

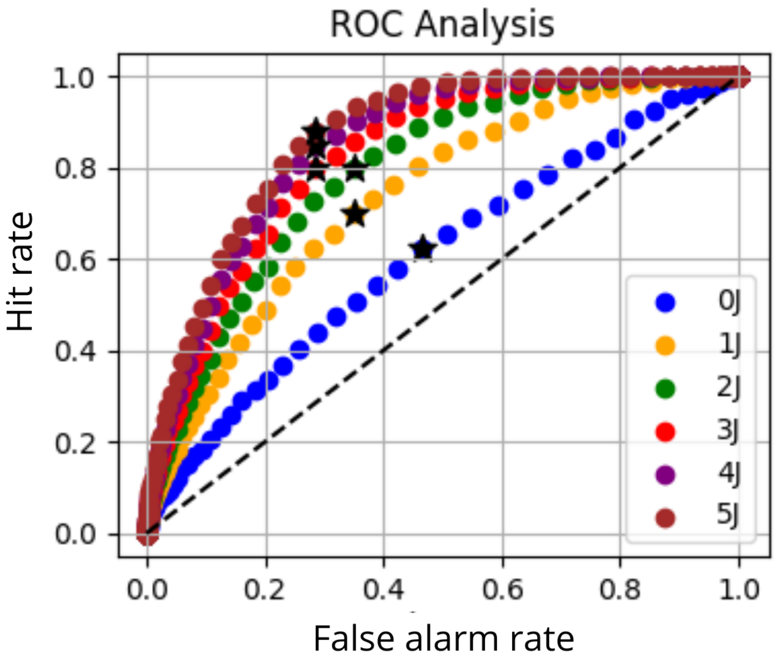
Skill of analog forecast with a temporal tolerance

n-day temporal tolerance

If an EPE is forecast between $J-n$ and $J+n$, is it observed between $J-n$ and $J+n$?

ROC analysis for Réunion :

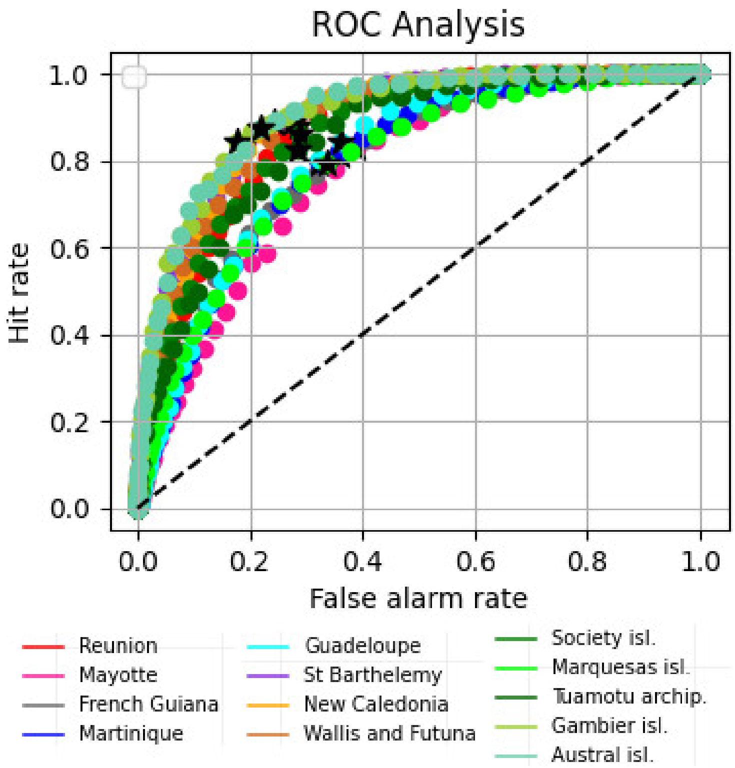
- The higher n is, the further the ROC curve is
- n = 5 days → Optimal $c_{Q,R}$ at H = 0.88, F = 0.29 ⇒ 75 % of EPE occurrence probability if an EPE is forecast



Tolerance ↗ = Predictability potential ↗
Definition of a **11-day window** with a higher probability of EPE occurrence
Perspectives for the subseasonal forecast, with a compromise between H, F, and the size of the window

Skill of analog forecast with a temporal tolerance for all OMs

And for all OMs ?



ROC analysis, 5-day tolerance :

- Predictability potential for all OMs
- Optimal $c_{Q,R}$: H = 0.80 to 0.90, F = 0.20 to 0.40

Tolerance ↗ = Predictability potential ↗
Definition of a **11-day window** with a higher probability of EPE occurrence

Conclusion and perspectives

General study of EPEs in all tropical French Overseas territories (**RR24** $\geq$ **p99**) on a long period of time and thanks to a rich rain gauge database.

Conclusion and perspectives

General study of EPEs in all tropical French Overseas territories (**RR24** $\geq$ **p99**) on a long period of time and thanks to a rich rain gauge database.

What is the average large-scale environment of EPEs ?

- **Large-scale strong humidity and convective activity** + cyclonic anomaly
- Main contribution : **intraseasonal**, especially with equatorial waves

EPE Weather Patterns

- Classification method, **common for all OMs**
- 6 EPEs weather patterns in Réunion, depending on the season and the large-scale structures. **Equatorial waves and interaction with midlatitudes**
- Preferred area of EPE occurrence depending on the WP = **impact of the large scale on the local scale**

Conclusion and perspectives

General study of EPEs in all tropical French Overseas territories (**RR24** $\geq$ **p99**) on a long period of time and thanks to a rich rain gauge database.

What is the average large-scale environment of EPEs ?

- **Large-scale strong humidity and convective activity** + cyclonic anomaly
- Main contribution : **intraseasonal**, especially with equatorial waves

EPE Weather Patterns

- Classification method, **common for all OMs**
- 6 EPEs weather patterns in Réunion, depending on the season and the large-scale structures. **Equatorial waves and interaction with midlatitudes**
- Preferred area of EPE occurrence depending on the WP = **impact of the large scale on the local scale**

Can the predictability of EPEs be increased thanks to large-scale indicators ?

Analog forecast

- Weather Patterns = EPE **Predictability potential** for all OMs
- **Temporal tolerance** $\nearrow$ **= potential** $\nearrow$

Conclusion and perspectives

General study of EPEs in all tropical French Overseas territories (**RR24 ≥ p99**) on a long period of time and thanks to a rich rain gauge database.

What is the average large-scale environment of EPEs ?

- **Large-scale strong humidity and convective activity** + cyclonic anomaly
- Main contribution : **intraseasonal**, especially with equatorial waves

EPE Weather Patterns

- Classification method, **common for all OMs**
- 6 EPEs weather patterns in Réunion, depending on the season and the large-scale structures. **Equatorial waves and interaction with midlatitudes**
- Preferred area of EPE occurrence depending on the WP = **impact of the large scale on the local scale**

Can the predictability of EPEs be increased thanks to large-scale indicators ?

Analog forecast

- Weather Patterns = EPE **Predictability potential** for all OMs
- **Temporal tolerance ↗ = potential ↗**

Thank you for your attention

Manuscript : https://meteofrance.hal.science/tel-05350912

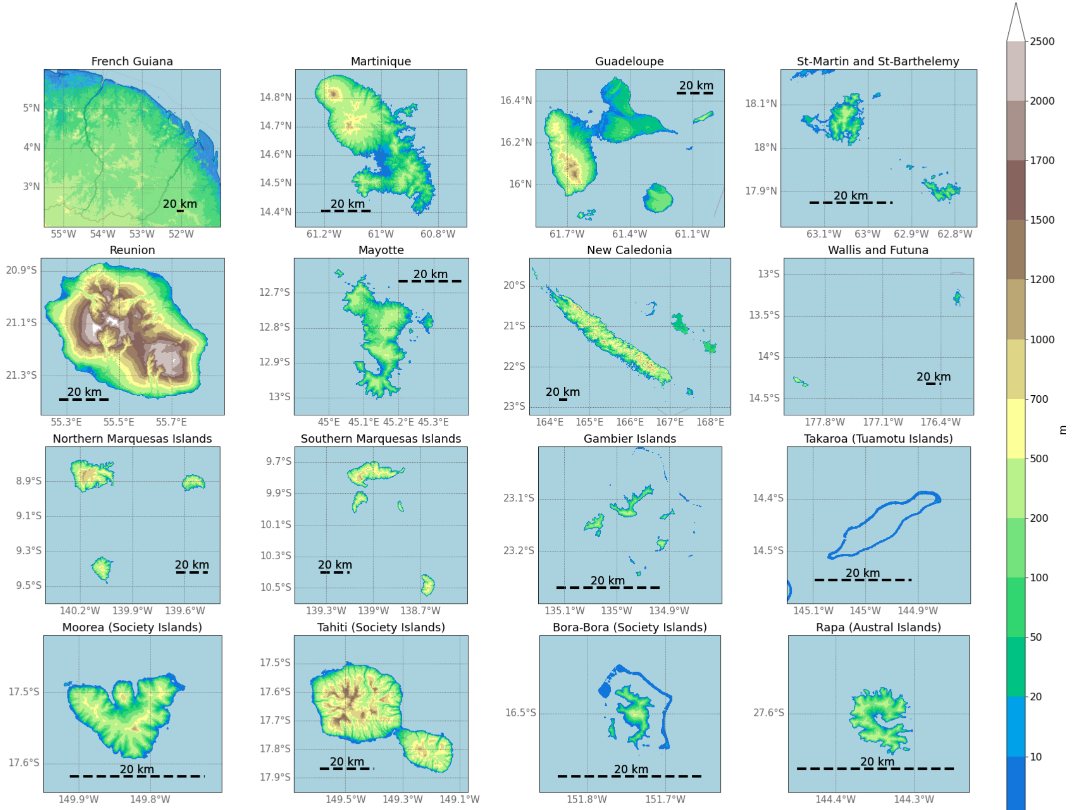
Appendix 1 : Rainfall data

Complex territories

- Diverse areas and topography, with many **tiny** islands (< 1000 km^2) and/or with a **sharp relief** (until 1500 to 3000 meters)
- Need for **high spatial resolution** rainfall data (local effects)
- Study of EPEs : **a large temporal size of data is essential**

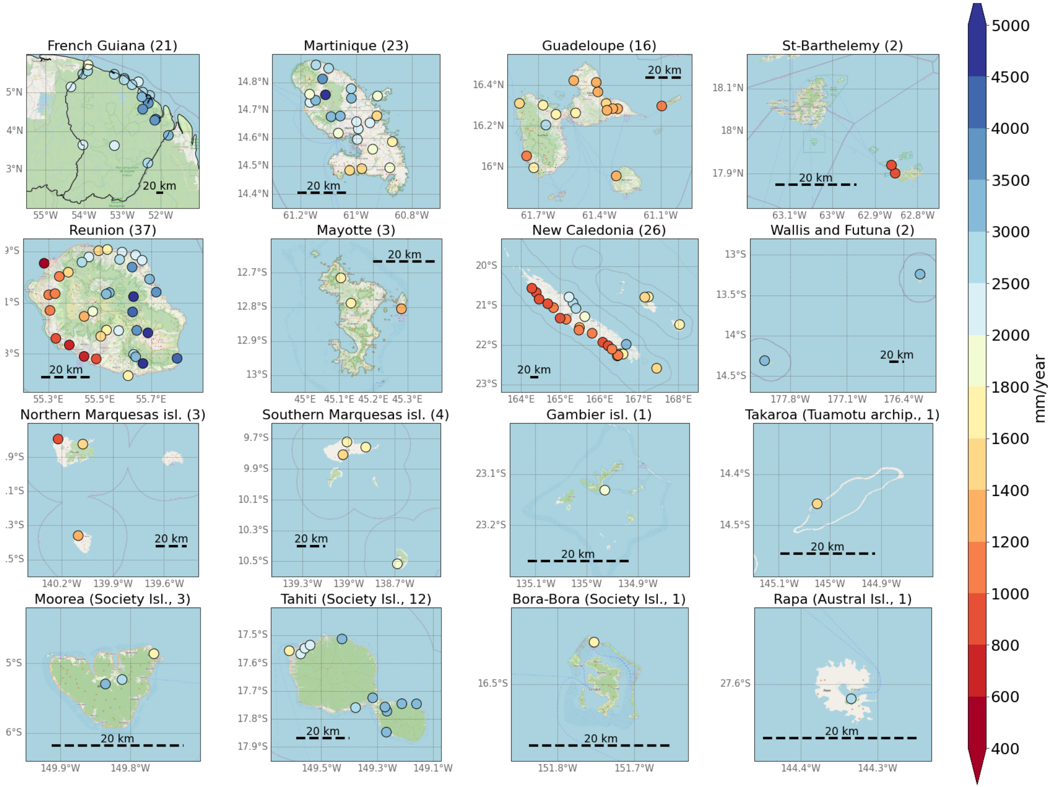
Satellite rainfall products

- Global and homogeneous coverage
- Poor estimation skill of rainfall over tiny islands (Rahmawati & Lubczynski, 2018) and during EPEs (Sanogo et al., 2022)



Topography of each OM

Appendix 1 : Rainfall data



Annual mean rainfall for each rain gauge

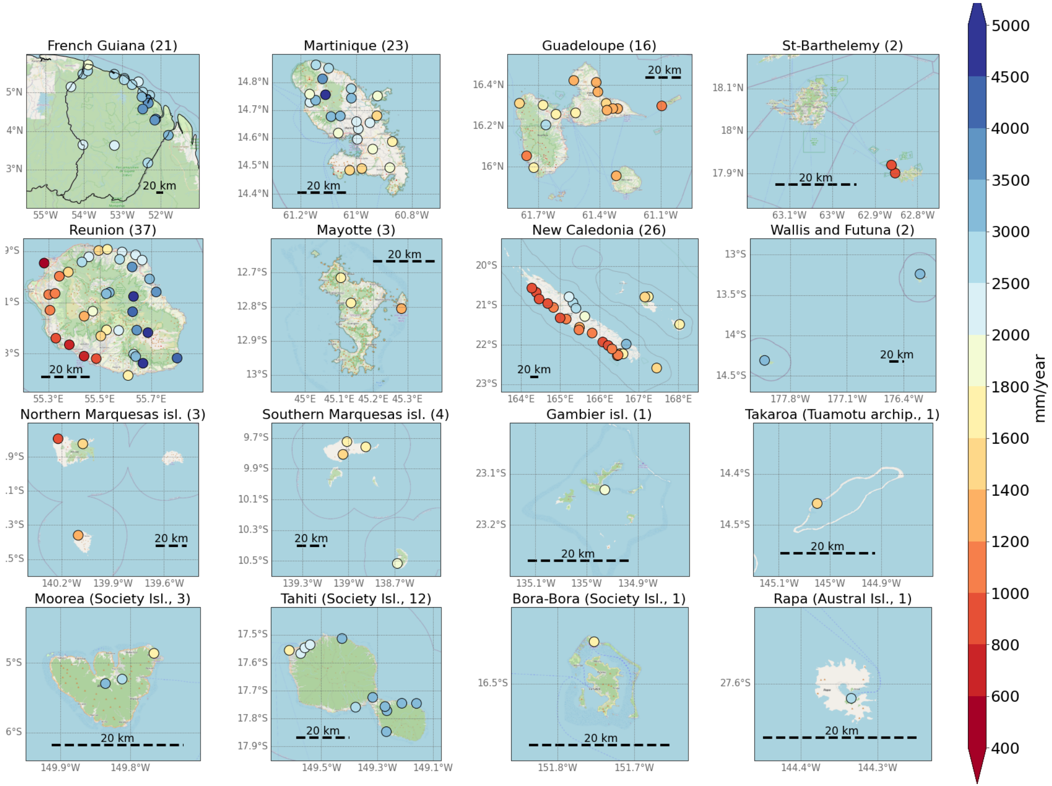
Météo-France network

- Dense network of rain gauges in many territories
- Large temporal size of data, especially for daily data
- Regularly expertised data

Daily rainfall : choice of rain gauges

- **1979-2021**
- Rain gauges with less than 5 % of missing data

Appendix 1 : Rainfall data



Annual mean rainfall for each rain gauge

Météo-France network

- Dense network of rain gauges in many territories
- Large temporal size of data, especially for daily data
- Regularly expertised data

Daily rainfall : choice of rain gauges

- **1979-2021**
- Rain gauges with less than 5 % of missing data

→ 156 rain gauges, some of which are above 1000 meters

High spatial variability of rainfall between territories (different climates) and within each territory (local effects)

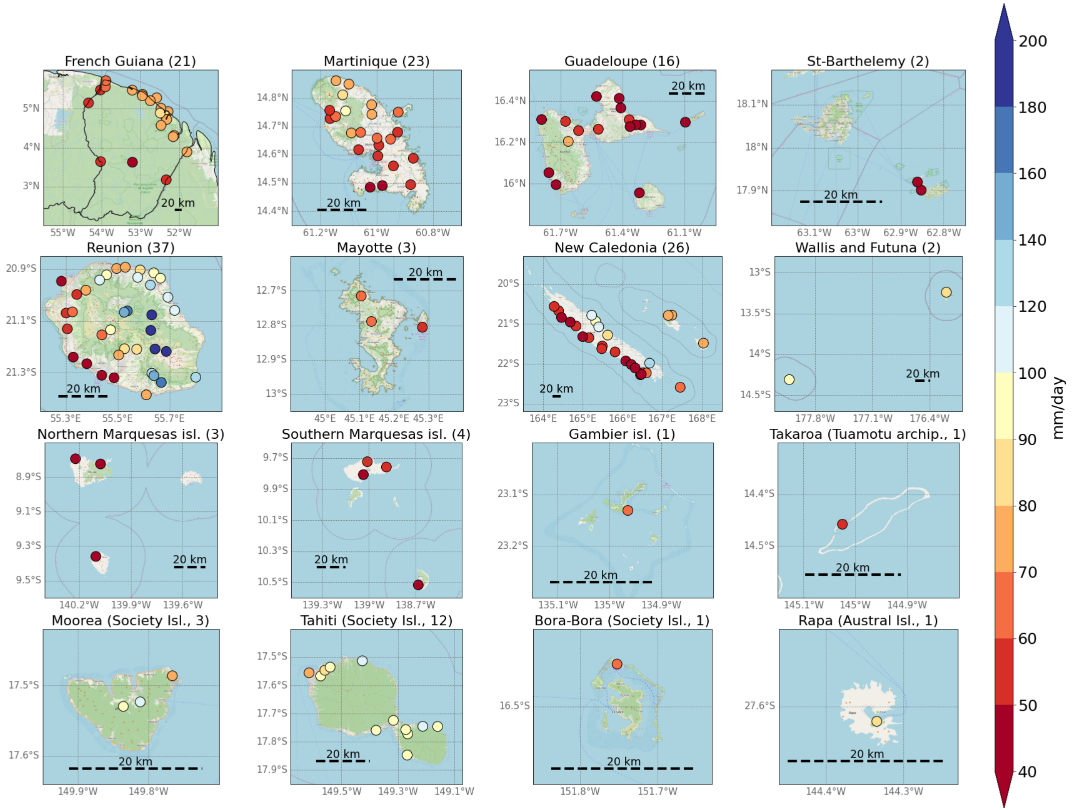
Appendix 2 : EPE definition

Définition d'un EPE

RR24 ≥ p99 (all-day) for each rain gauge

→ ≃ **150 to 160 EPEs** for each station

Different thresholds depending on the rain gauge to **take into account the high variability of rainfall climatology** (*inter-OM* and *intra-OM* variability)



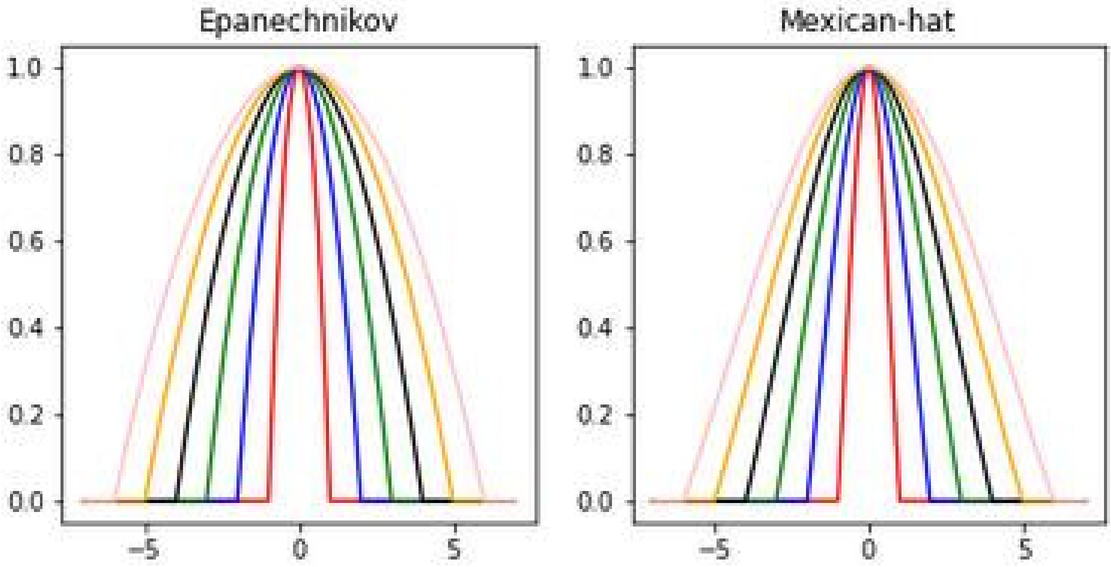
P99 for each rain gauge

Appendix 3 : Self-Organizing Maps

The lateral interaction between neighbouring nodes is dealt with a **neighbouring function**, depending on a **neighbouring radius** which evolve with radius. Higher in the first iterations, it has to decrease to insure the map convergence (Kohonen, 1982)

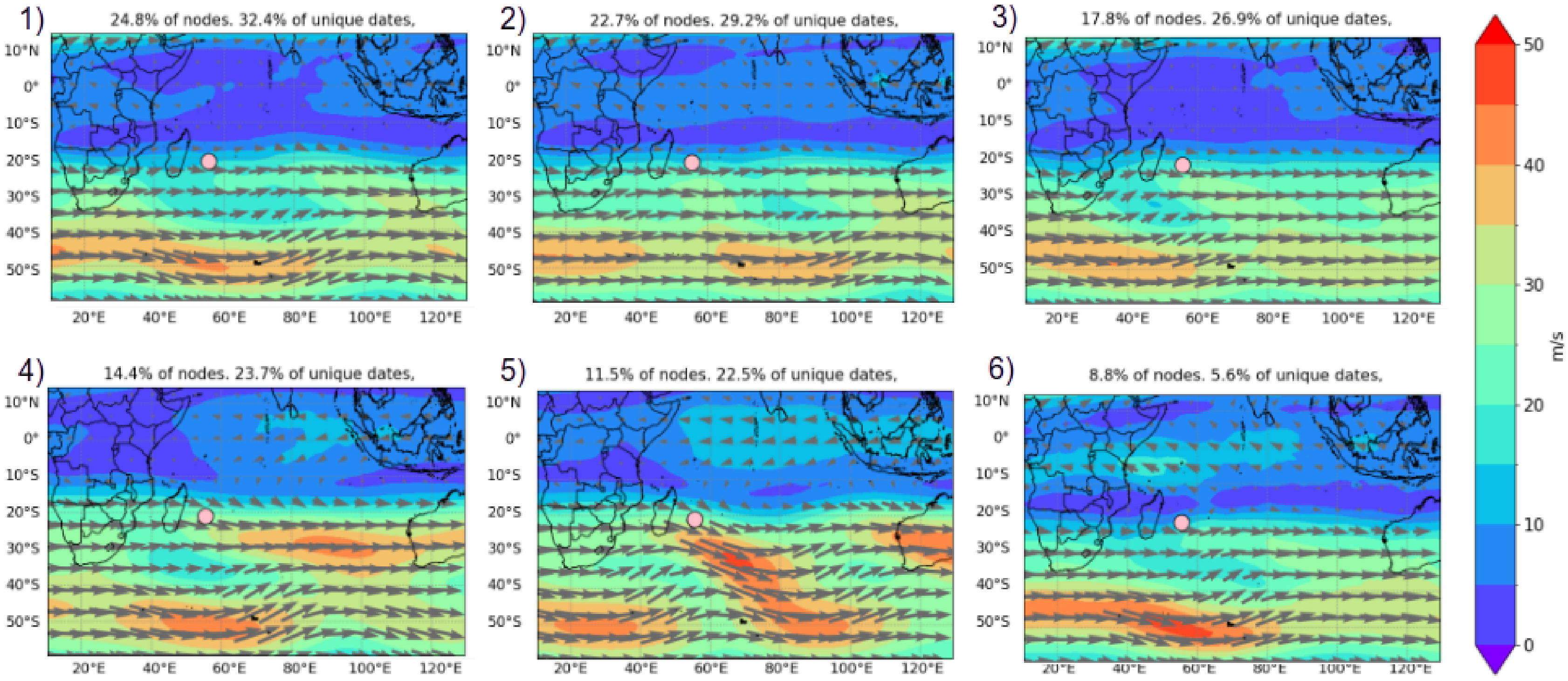
- too slow decrease : too many lateral interactions, which lead to an homogeneous map (all nodes are equals)
- too fast decrease : lack of interactions between neighbouring nodes, all data can be assigned to a reduce number of nodes

The neighbouring function gives the ponderation applied when other nodes than the BMU is updated, depending on the distance on the map between the BMU and other nodes.



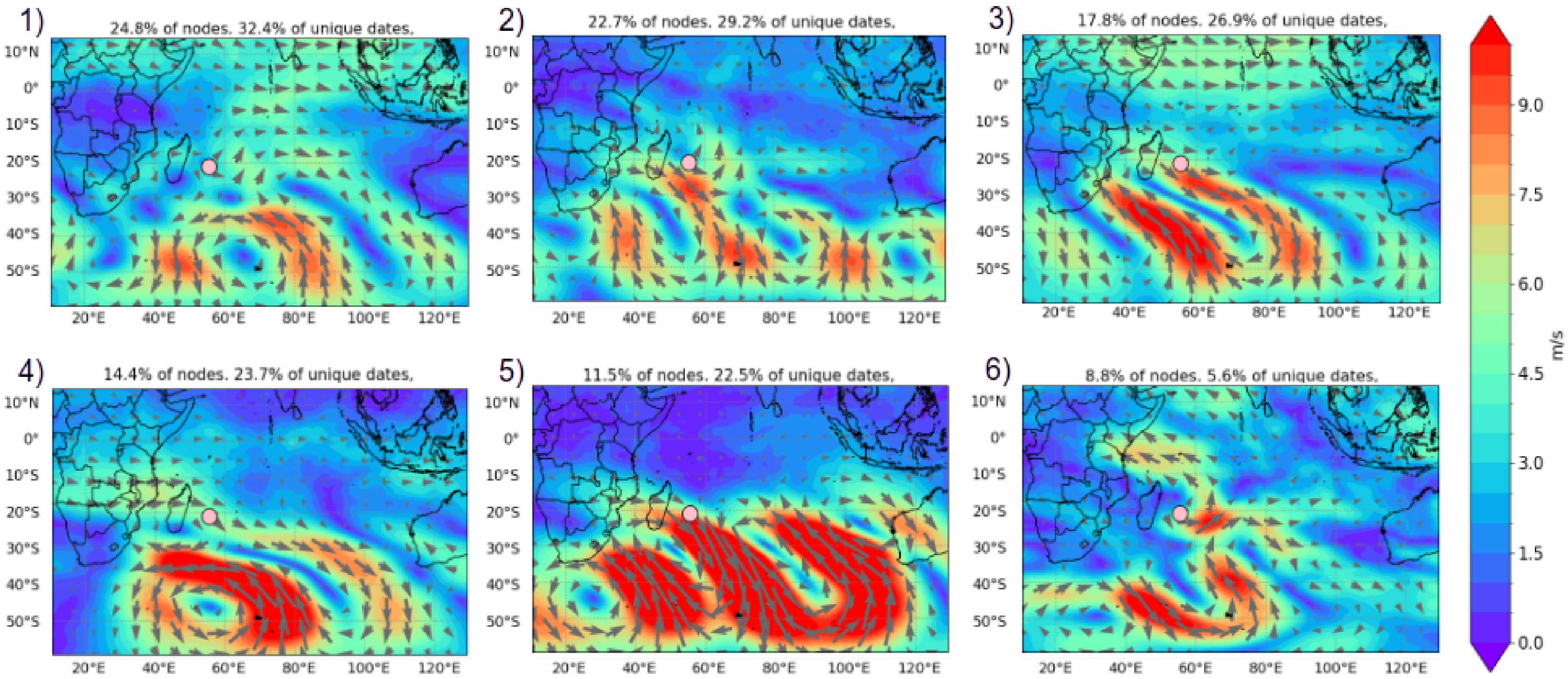
Neighbouring functions disponibles in SuSi, with neighbouring radius from 6 to 1.

Appendix 4 : WPs in la Réunion



200 hPa wind composite for each WP

Appendix 4 : WPs in la Réunion

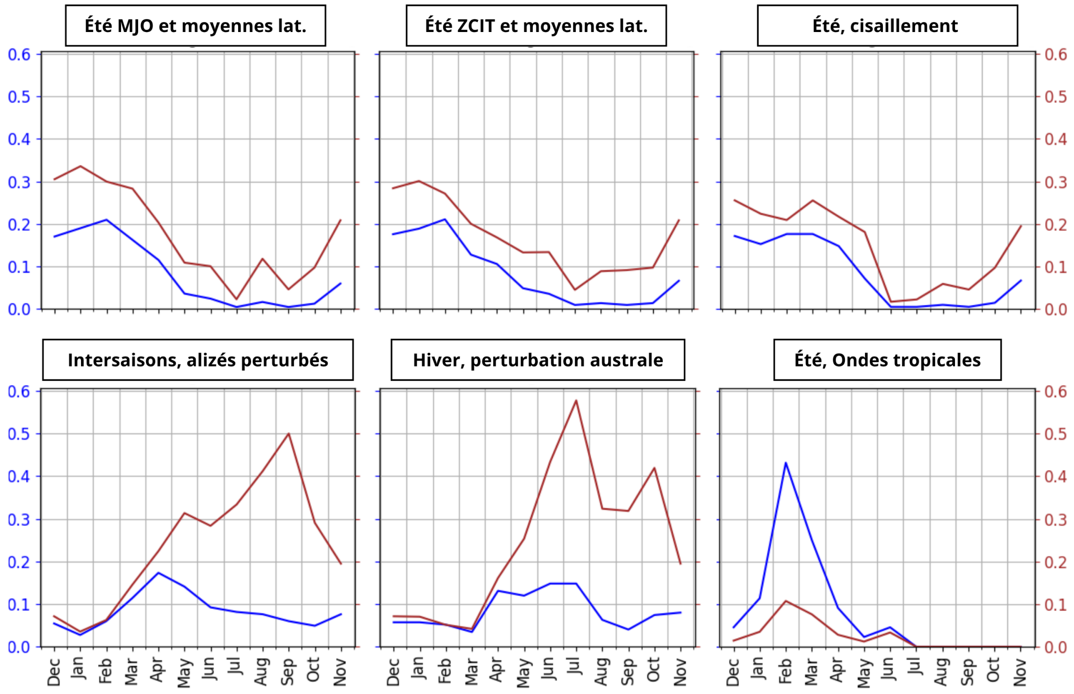


200 hPa wind anomaly composite for each WP

Appendix 5 : Annual cycle of EPE weather patterns in Réunion

Fraction of monthly occurrence of WPs, relative to each WP et relative to each month

Consistency between characteristics of WPs and their preferred occurrence period
But, only with the last WP, possible occurrence of WPs all year long

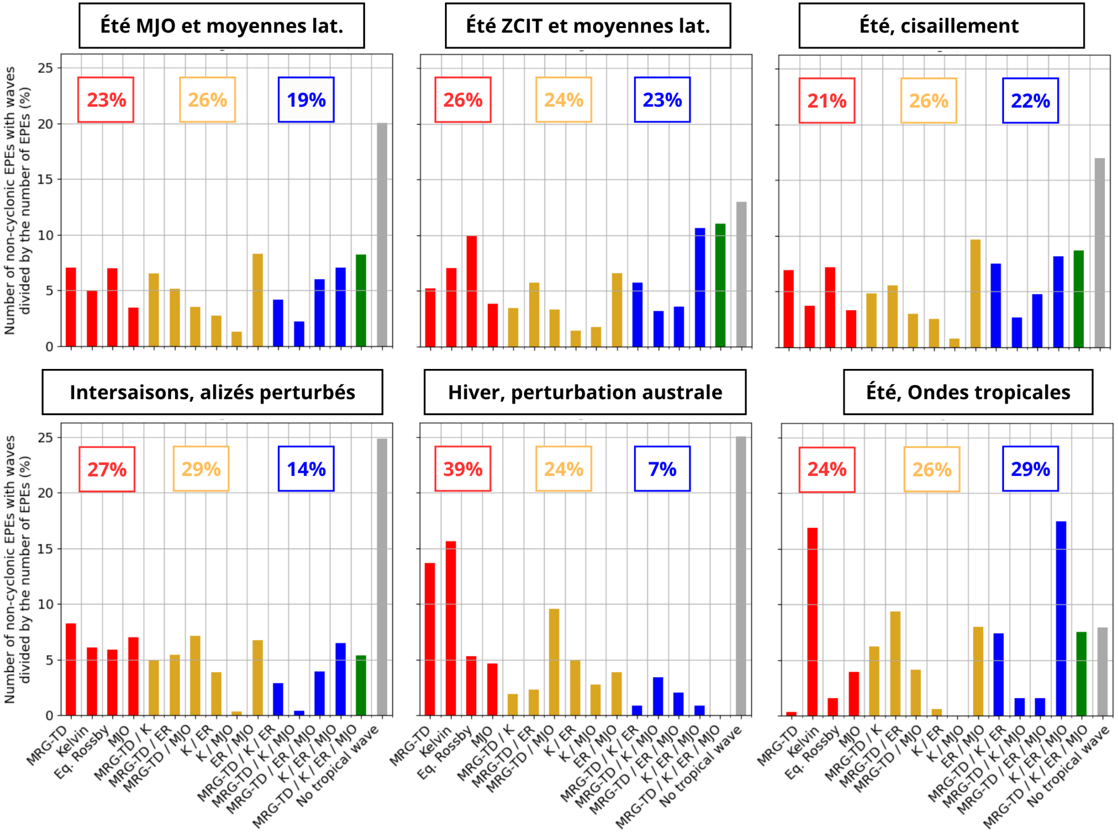


Appendix 6 : Tropical variability for each WP

Do specific wave configurations emerge in each WP ?

- Less days without active wave for the summer WPs (8 to 20%) than for shoulder seasons and winter WPs (25%)
- Configuration with 3 or 4 active waves more frequent during summer WPs
- Austral disturbances : more frequent with configurations with 1 active wave (Kelvin wave = projection of midlatitudes Rossby waves)

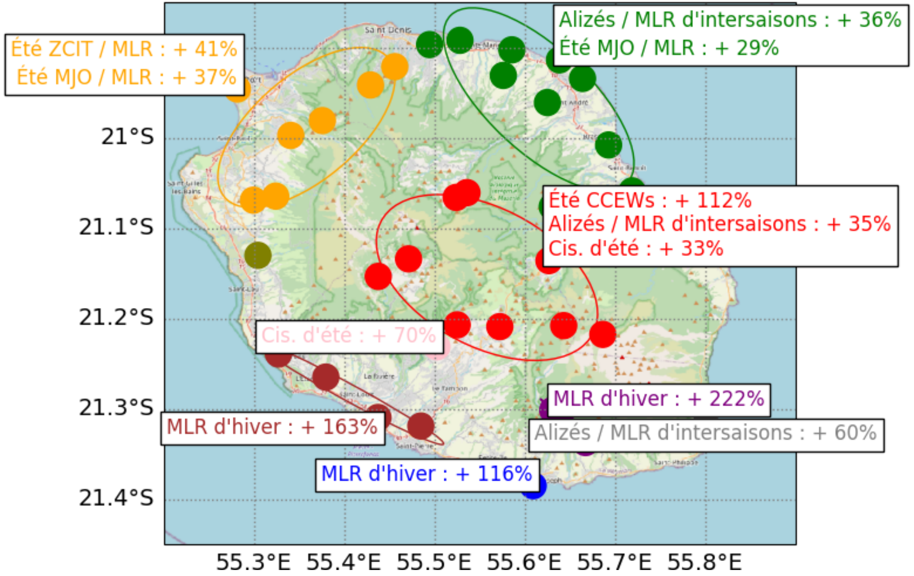
Fractions of dates of each WP with active waves like the configuration described.



Appendix 7 : Preferred EPE occurrences depending on the area

Agglomerative Hierarchical clustering of rain gauges according to their common EPE dates → **distinct geographical areas with common EPE dates**

Are EPE from a rain gauge cluster associated with preferred WPs ?



EPE Probability modulation of rain gauge compared to the rest of the island :

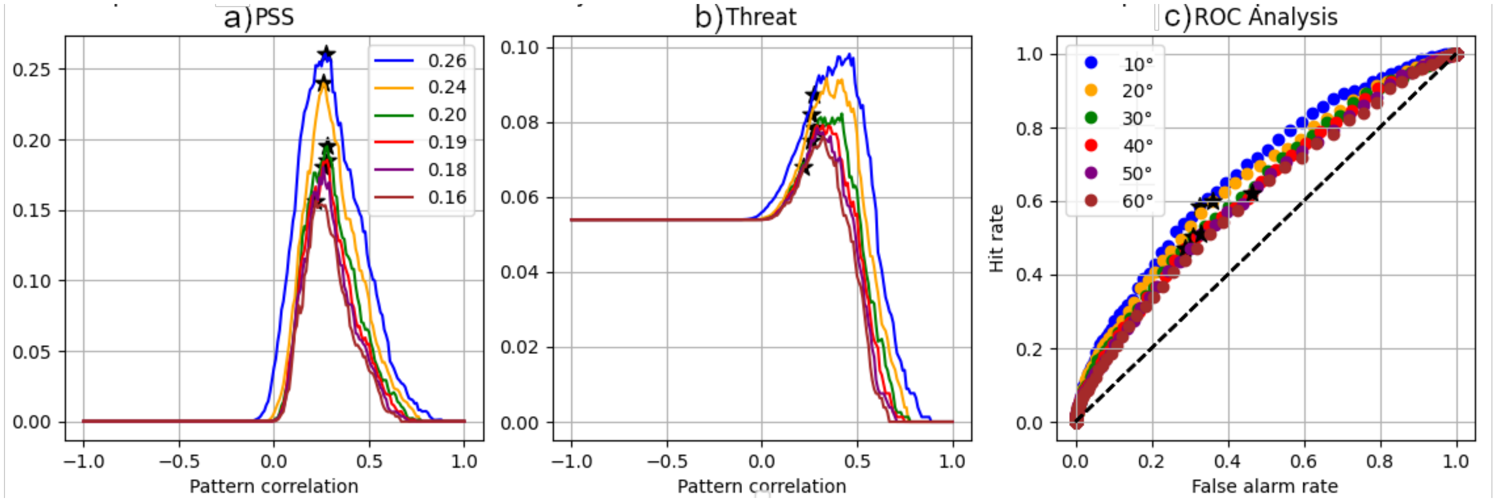
1) Été MJO / MLR : north
2) Été ZCIT / MLR : northwest
3) Cis. d'été : centre / south
4) Alizés / MLR intersaison : east / centre
5) MLR d'hiver : south
6) Été CCEWs : centre

Preferred weather patterns = **consistency between the low-level wind anomaly (humidity flux) and the exposition of rain gauges regarding of topography**

Appendix 8 : Forecast scores depending on the geographic area

Evaluation of potential predictability of EPEs$_{nc}$ occurrence :

- PSS (*Pierce Skill Score*) : *hit rate* (H) - *False alarm* (F)
- Threat : number of good forecasts of occurrence / observed and forecast occurrence
- ROC : H as a function of F



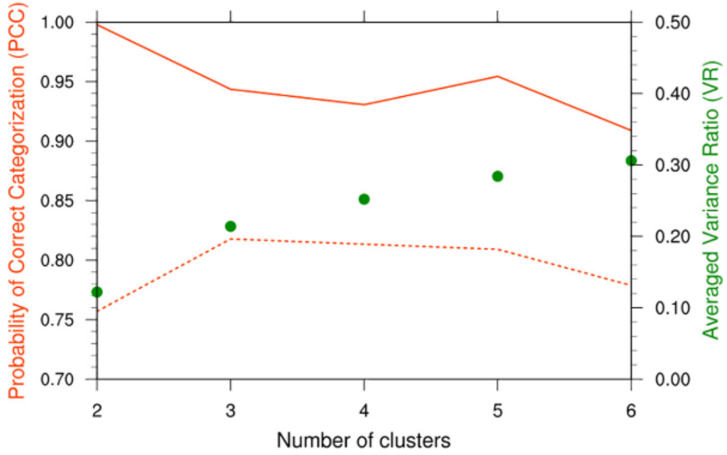
Forecast scores sensitivity to the size of the correlation area, as a function of the threshold of spatial correlation

- Over the training area, low potential of predictability (PSS max = 0.16) for the day
- Correlation size area ↘ = PSS ↗

Appendix 9 : SOM + k-means classification vs. simple k-means ?

Wang et al. (2022) : clustering analysis performed on the 850 hPa large-scale flow fields centred at the tropical cyclone (TC) genesis, to determine the cyclogenesis conditions in the Pacific Northwest Ocean. 5 weather patterns are identified, comparable between SOMs and k-means.

Circulation patterns	SOM	K-means
Convergence (CON)	125 (27.1%)	109 (23.6%)
Monsoon trough (MT)	114 (24.7%)	87 (18.9%)
Pre-existing cyclone (PC)	81 (17.6%)	82 (17.8%)
Subtropical high (SH)	74 (16.1%)	87 (18.9%)
Easterly wave (EW)	67 (14.5%)	96 (20.8%)

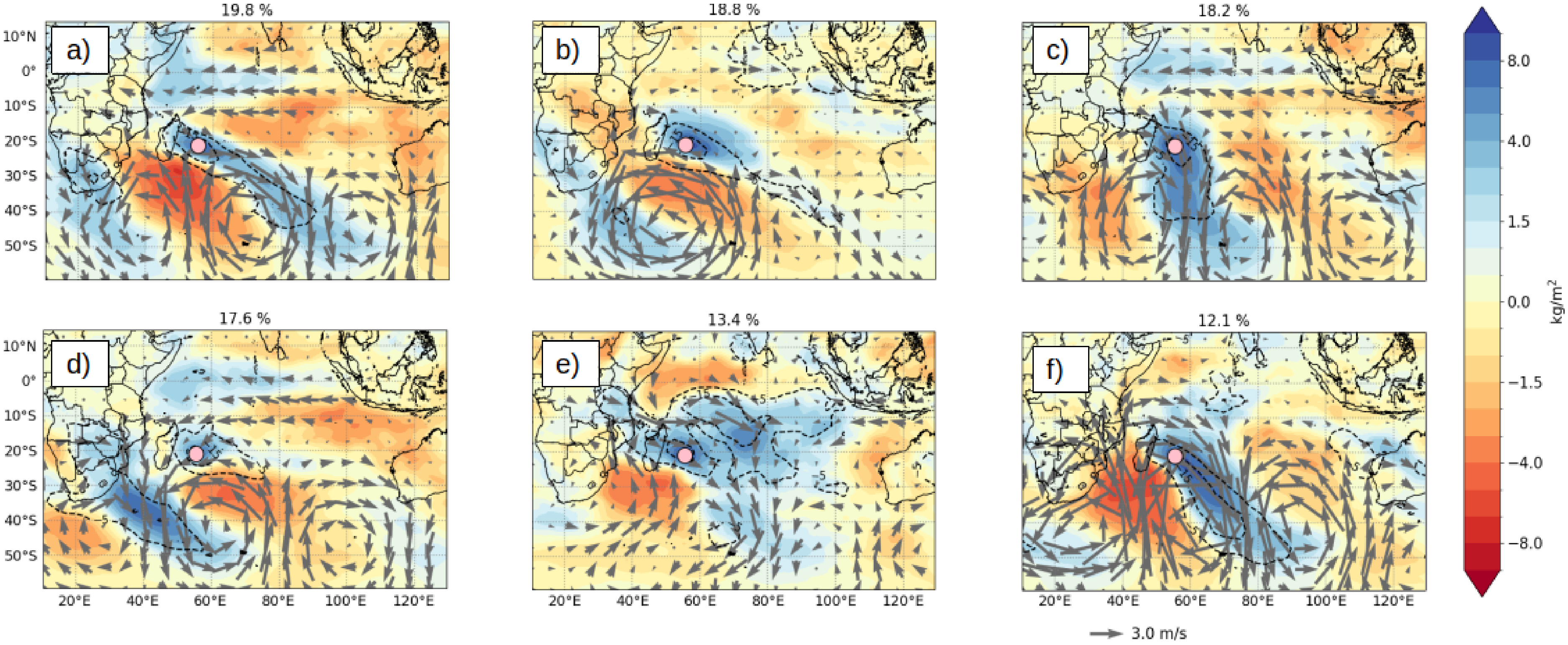


Left : Numbers and percentage (parenthesis) of cases in 5 WPs derived from the SOM and k-means clustering, respectively. Right : Probability of correct categorization (PCC, SOM : red line, k-means : red dots) and averaged variance ratio (VR, green, only for SOMs) corresponding to varying number of clusters.

- whatever the final number of WPs, **better PCC for SOM**
- **k-means tend to form clusters with homogeneous size** $\longrightarrow$ better discrimination between different situations with SOM

Appendix 9 : SOM + k-means classification vs. simple k-means ?

Weather patterns with a direct k-means classification of $EPEs_{nc}$ dates, without SOM



Appendix 9 : SOM + k-means classification vs. simple k-means ?

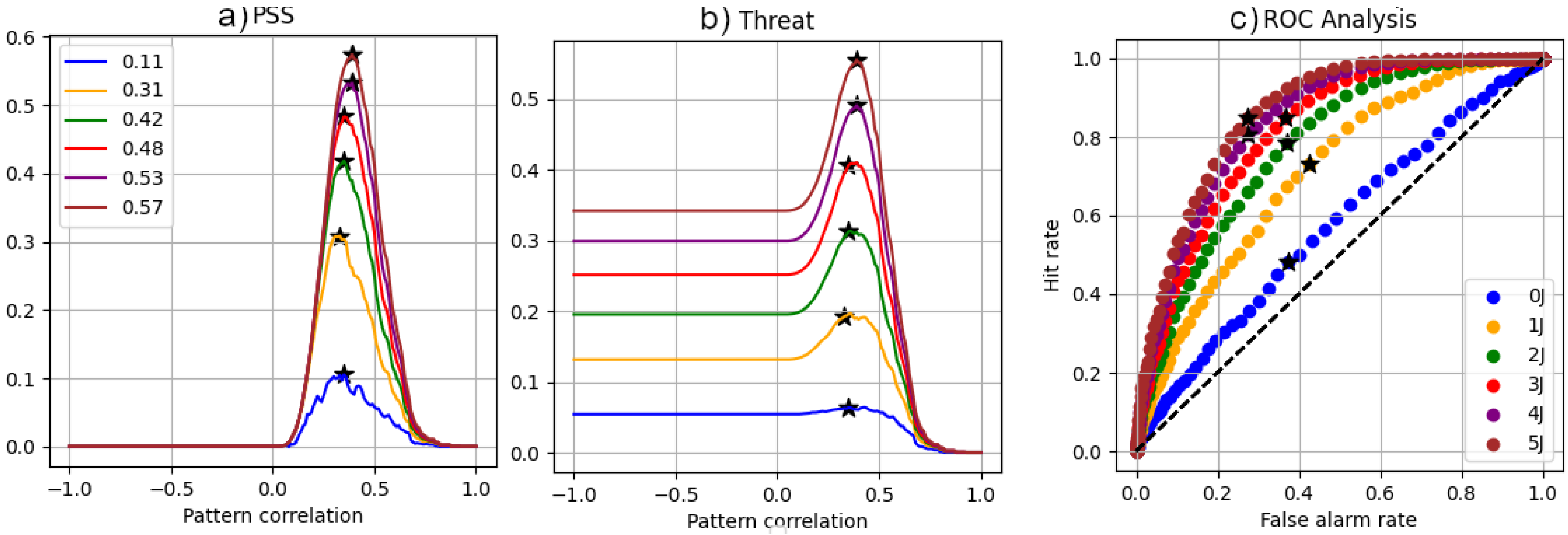
Weather patterns with a direct k-means classification of EPEs$_{nc}$ dates, without SOM

WPs	WP 1 k-means (19,8 %)	WP 2 k-means (18,8 %)	WP 3 k-means (18,2 %)	WP 4 k-means (17,6 %)	WP 5 k-means (13,4 %)	WP 6 k-means (12,1 %)
WP 1 SOM + k-means (32.4 %)	-0.10	-0.50	**0.66**	**0.66**	-0.06	-0.59
WP 2 SOM + k-means (29.2 %)	-0.01	-0.37	**0.32**	-0.37	**0.67**	0.13
WP 3 SOM + k-means (26.9 %)	**0.65**	0.01	-0.39	**0.34**	-0.12	-0.48
WP 4 SOM + k-means (23.7 %)	-0.52	**0.62**	0.11	**0.35**	-0.38	-0.32
WP 5 SOM + k-means (22.5 %)	0.13	**0.30**	-0.53	-0.79	0.05	**0.89**
WP 6 SOM + k-means (5.6 %)	-0.34	-0.20	0.21	0.18	**0.69**	-0.43

Spatial correlations between SOMs + k-means WPs and only k-means WPs. Correlations > 0.29 are bolded.

Appendix 9 : SOM + k-means classification vs. simple k-means ?

Evaluation of predictability of EPE_{nc} occurrence



- Less skillful scores : slightly worse classification of $EPEs_{nc}$ situations
- Higher optimal spatial correlation threshold
- Correlation size area ↗ = PSS ↗ (≠ SOMs + k-means)
- Temporal tolerance = improvements of scores

Cornillault, E., Peyrille, P., Couvreux, F., & Roehrig, R. (2024).Large-scale drivers of tropical extreme precipitation events : The example of french overseas territories [_eprint : https://onlinelibrary.wiley.com/doi/pdf/10.1029/2024GL108770]. *Geophysical Research Letters*, *51*(15), e2024GL108770. https://doi.org/10.1029/2024GL108770

CRED & UNDRR. (2020, octobre 12). *The human cost of disasters : An overview of the last 20 years (2000-2019)*. Centre for Research on the Epidemiology of Disasters (CRED). Brussels, Belgium. Récupérée octobre 4, 2023, à partir de https://reliefweb.int/report/world/human-cost-disasters-overview-last-20-years-2000-2019

Grotjahn, R., Black, R., Leung, R., Wehner, M. F., Barlow, M., Bosilovich, M., Gershunov, A., Gutowski, W. J., Gyakum, J. R., Katz, R. W., Lee, Y.-Y., Lim, Y.-K., & Prabhat. (2016).North american extreme temperature events and related large scale meteorological patterns : A review of statistical methods, dynamics, modeling, and trends. *Climate Dynamics*, *46*(3), 1151-1184. https://doi.org/10.1007/s00382-015-2638-6

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020).The ERA5 global reanalysis [dataset] [_eprint : https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803]. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. https://doi.org/10.1002/qj.3803

Kohonen, T. (1982).Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59-69. https://doi.org/10.1007/BF00337288

Latos, B., Lefort, T., Flatau, M. K., Flatau, P. J., Permana, D. S., Baranowski, D. B., Paski, J. A. I., Makmur, E., Sulystyo, E., Peyrillé, P., Feng, Z., Matthews, A. J., & Schmidt, J. M. (2021).Equatorial Waves Triggering Extreme Rainfall and Floods in Southwest Sulawesi, Indonesia [Publisher : American Meteorological Society Section : Monthly Weather Review]. *Monthly Weather Review*, *149*(5), 1381-1401. https://doi.org/10.1175/MWR-D-20-0262.1

Liebmann, B., & Smith, C. (1996).Description of a complete (interpolated) outgoing longwave radiation dataset. *Bulletin of the American Meteorological Society*, *77*(6), 1275-1277. https://psl.noaa.gov/data/gridded/OLRinterpolated.pdf

Lorenz, E. N. (1969).Atmospheric Predictability as Revealed by Naturally Occurring Analogues [Publisher : American Meteorological Society Section : Journal of the Atmospheric Sciences]. *Journal of the Atmospheric Sciences*, *26*(4), 636-646. https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2

Moron, V., Gouirand, I., & Taylor, M. (2016).Weather types across the caribbean basin and their relationship with rainfall and sea surface temperature. *Climate Dynamics*, *47*(1), 601-621. https://doi.org/10.1007/s00382-015-2858-9

Moron, V., Robertson, A. W., Qian, J.-H., & Ghil, M. (2015).Weather types across the Maritime Continent : from the diurnal cycle to interannual variations [Publisher : Frontiers]. *Frontiers in Environmental Science*, *2*. https://doi.org/10.3389/fenvs.2014.00065

Nguyen-Le, D., & Yamada, T. J. (2019). Using weather pattern recognition to classify and predict summertime heavy rainfall occurrence over the upper nan river basin, northwestern thailand (supporting information). *Weather and Forecasting*, *34*(2), 345-360. https://doi.org/10.1175/WAF-D-18-0122.1

Nguyen-Le, D., Yamada, T. J., & Tran-Anh, D. (2017). Classification and forecast of heavy rainfall in northern kyushu during baiu season using weather pattern recognition [_eprint : https://onlinelibrary.wiley.com/doi/pdf/10.1002/asl.759]. *Atmospheric Science Letters*, *18*(8), 324-329. https://doi.org/10.1002/asl.759

Peyrillé, P., Roehrig, R., & Sanogo, S. (2023). Tropical waves are key drivers of extreme precipitation events in the central sahel [_eprint : https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL103715]. *Geophysical Research Letters*, *50*(20), e2023GL103715. https://doi.org/10.1029/2023GL103715

Rahmawati, N., & Lubczynski, M. W. (2018). Validation of satellite daily rainfall estimates in complex terrain of bali island, indonesia. *Theoretical and Applied Climatology*, *134*(1), 513-532. https://doi.org/10.1007/s00704-017-2290-7

Rapolaki, R. S., Blamey, R. C., Hermes, J. C., & Reason, C. J. C. (2019). A classification of synoptic weather patterns linked to extreme rainfall over the limpopo river basin in southern africa. *Climate Dynamics*, *53*(3), 2265-2279. https://doi.org/10.1007/s00382-019-04829-7

Réchou, A., Flores, O., Jumaux, G., Duflot, V., Bousquet, O., Pouppeville, C., & Bonnardot, F. (2019). Spatio�temporal variability of rainfall in a high tropical island : Patterns and large�scale drivers in réunion island. *Quarterly Journal of the Royal Meteorological Society*, *145*(720), 893-909. https://doi.org/10.1002/qj.3485

Sanogo, S., Peyrillé, P., Roehrig, R., Guichard, F., & Ouedraogo, O. (2022). Extreme Precipitating Events in Satellite and Rain Gauge Products over the Sahel [Publisher : American Meteorological Society Section : Journal of Climate]. *Journal of Climate*, *35*(6), 1915-1938. https://doi.org/10.1175/JCLI-D-21-0390.1

Wang, Z.-Q., Chen, G.-H., & Zhao, Y.-H. (2022). Clustering of large-scale circulation patterns during tropical cyclone preformation period using self-organizing maps [_eprint : https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.7626]. *International Journal of Climatology*, *42*(13), 7013-7027. https://doi.org/10.1002/joc.7626