

Un Générateur stochastique de Temps pour l'Innovation (GTI)

Pascal Yiou (LSCE)

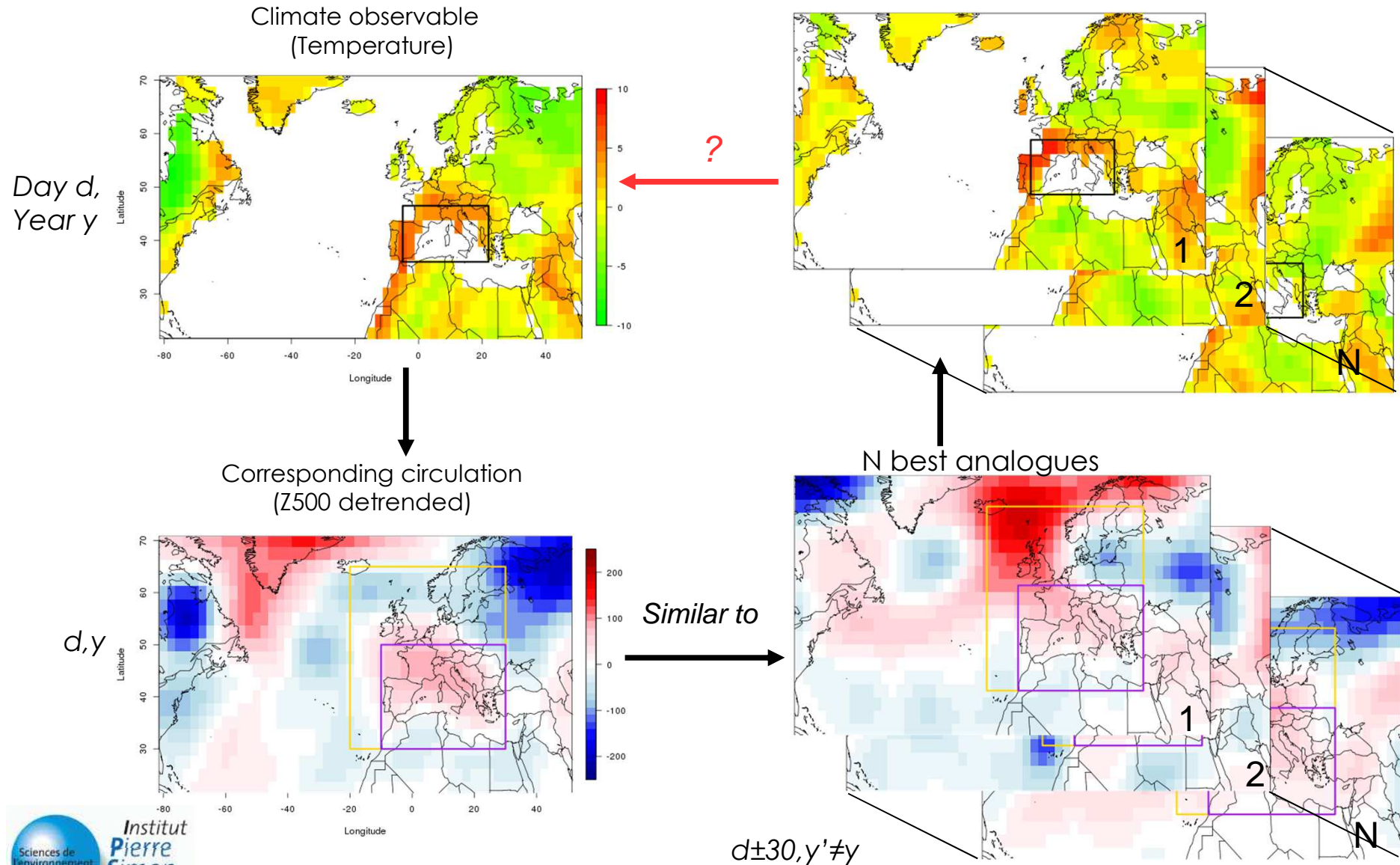
Céline Déandréis (ARIA)

Le générateur de temps AnaWEGE

- Développé dans le cadre de 2 projets amonts:
 - E3P (Climate KIC): générateur de temps stochastique & vagues de froid
 - A2C2 (ERC): Analogues de circulation
- Objectif: simuler de nombreuses et longues séries de variables climatiques présentant des propriétés physiques et climatiques plausibles
 - Générer de larges ensembles de saison (90 jours) à partir de conditions initiales spécifiées ou non.
 - Générer de longues séquences en climat stationnaire
- Principe: méthode des analogues
- Outil validé sur les températures:
 - Simulation des séries de température de 291 stations ECA&D
 - Analogues de circulation générale: pression au niveau de la mer
- Publication scientifique dans Geophysics Model Development
 - « AnaWEGE: a weather generator based on analogues of atmospheric circulation », P. Yiou, 2013



Circulation analogues





Le projet GTI



Objectifs

- Adapter le mode scénario de AnaWEGE en mode prédictif
- Comparaison de la performance de l'outil pour plusieurs variables:
 - Température de surface et précipitation
- Définir les régions et horizons temporels pour lesquels l'outil présente de bonnes performances
 - Simulation en mode "hindcast", comparaison avec les données disponibles: analyses & observations
 - Test pour la prévision court-terme et la prévision saisonnière
- Analyser l'intérêt de cet outil sur les marchés et produits d'ARIA-technologies
 - Intégration dans des produits existants
 - Produit en cours de développement
 - Produit à venir

- LSCE (coordination: P. Yiou)
 - Expertise sur l’outil AnaWEGE,
 - Transformation du modèle en mode prévision
 - Co-encadrement des stagiaires
- ARIA Technologies (coordination: C. Déandreis)
 - Réalisation les tests et évaluation des performances de l’outil sous la supervision de P. Yiou
 - Développement des modules de visualisation et de l’automatisation
 - Co-encadrement des stagiaires
 - Définition et analyse du marché potentiel

Statut du projet

- Les difficultés à la signature du contrat ont retardé le projet
- Réalisation de 2 stages:
 - Mariette Lamige (M2 U Lyon, Avril-September 2015)
 - Zhongya Liu (M1 X, Avril-Aout 2017)
- Reliquat d'argent pour poursuivre le travail via le recrutement d'un 3^{ème} stagiaire en 2018.
- Principales difficultés
 - Recrutement de stagiaires
 - L'accès à des données pour valider l'outil et le comparer à d'autres systèmes de prévision

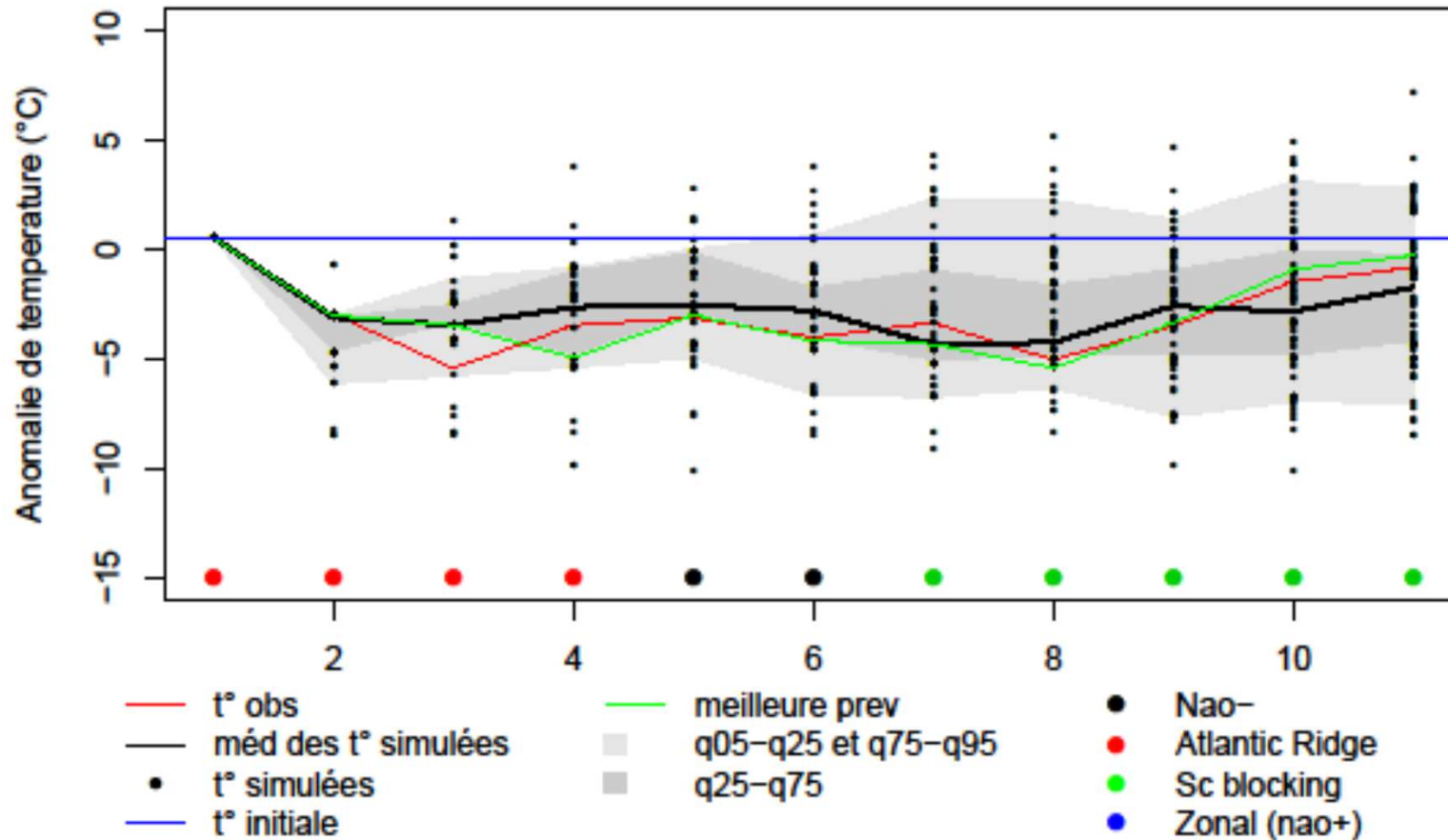
Avancées techniques

Tâches réalisées

- Implémentation du mode prévision et de l'automatisation
 - Mise en place d'une prévision hebdomadaire de la température sur 2 villes
- Implémentation des scores de performance et des modules de visualisation
- Validation du mode prévision pour la température en mode « hindcast » sur 2 villes:
 - Réalisation de simulations en mode "hindcast" sur la période de 2010 à aujourd'hui:
 - Simulations d'une durée de 5 jours à 90 jours
 - Lancement tous les 5 jours
 - Réalisation d'un ensemble de 100 simulations
 - À partir des ré-analyses NCEP (1948 à aujourd'hui) et des données aux stations ECA&D
 - Evaluation:
 - Définition de scores basées sur la RMSE,
 - Utilisation de scores classiques en prévision d'ensemble: le CRPS (Continuous Rank Probability Score), et le CRPSS (Continuous Rank Probability Skill Score)
 - Comparaison par rapport à la climatologie, à la persistance, aux observations

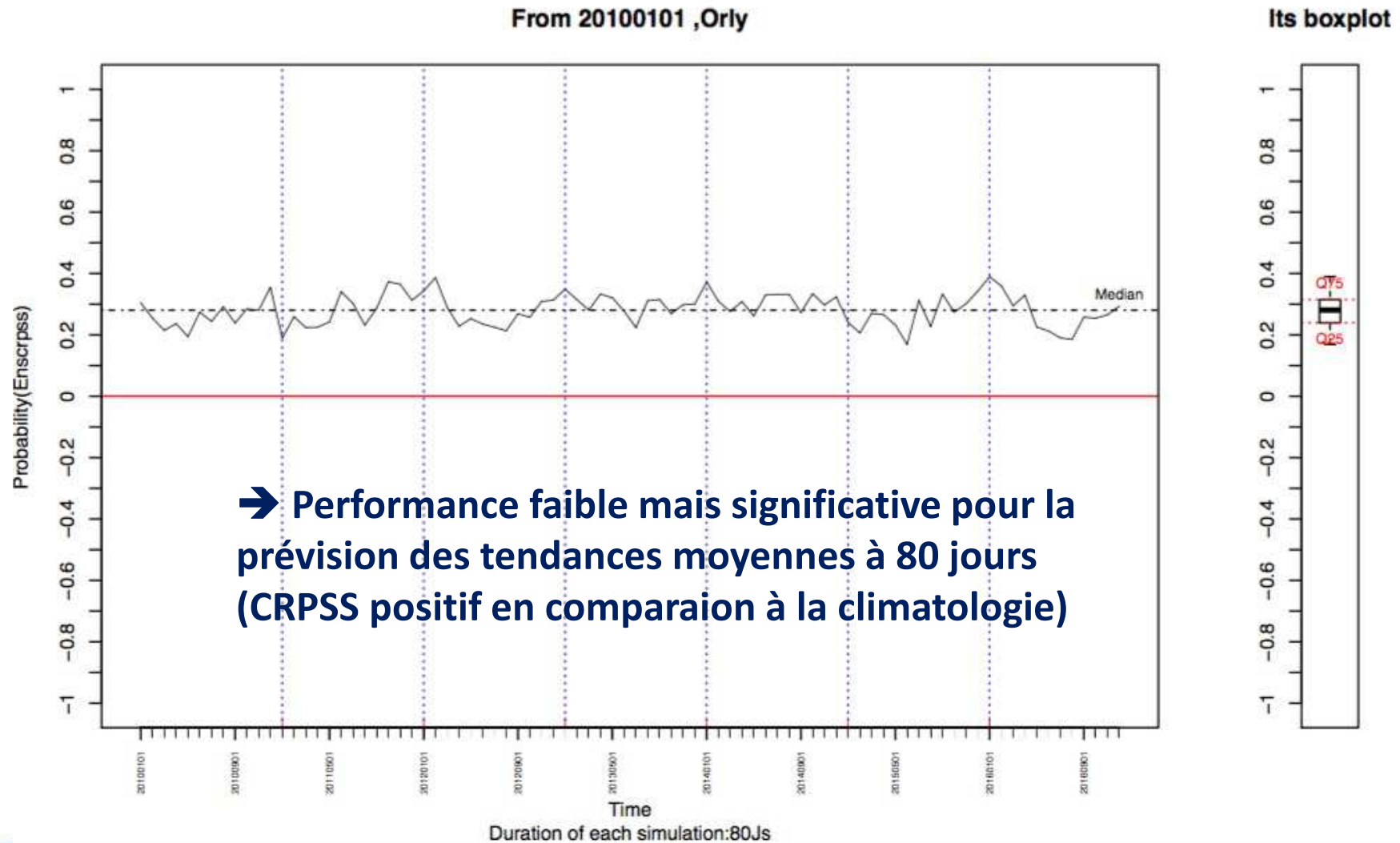
Stage M. Lamige (2015)

Paris – Jours depuis le 23/2/2001



➔ Performance à 5-6 jours pour les variables journalières

Performance à 80 jours



➔ Performance faible mais significative pour la prévision des tendances moyennes à 80 jours (CRPSS positif en comparaison à la climatologie)

Forces et faiblesses

- Forces:
 - Outil potentiellement utilisable pour la génération de données multi-sites, multivariés
 - Production de très grande base de données
 - Temps de calcul faible: 100 membres tournent sur un PC
- Faiblesses:
 - Nécessite de grande base de données: prédicteurs+variables à prédire
 - ➔ Coût élevé pour des utilisations commerciales
 - Ne permet pas de produire des évènements non observés

Intégration dans l'offre d'ARIA

3 axes identifiés

1. Développement de systèmes d'alerte pour les évènements de précipitations extrêmes (projet WEC)
2. Prévision saisonnière :
3. Nowcasting de précipitation

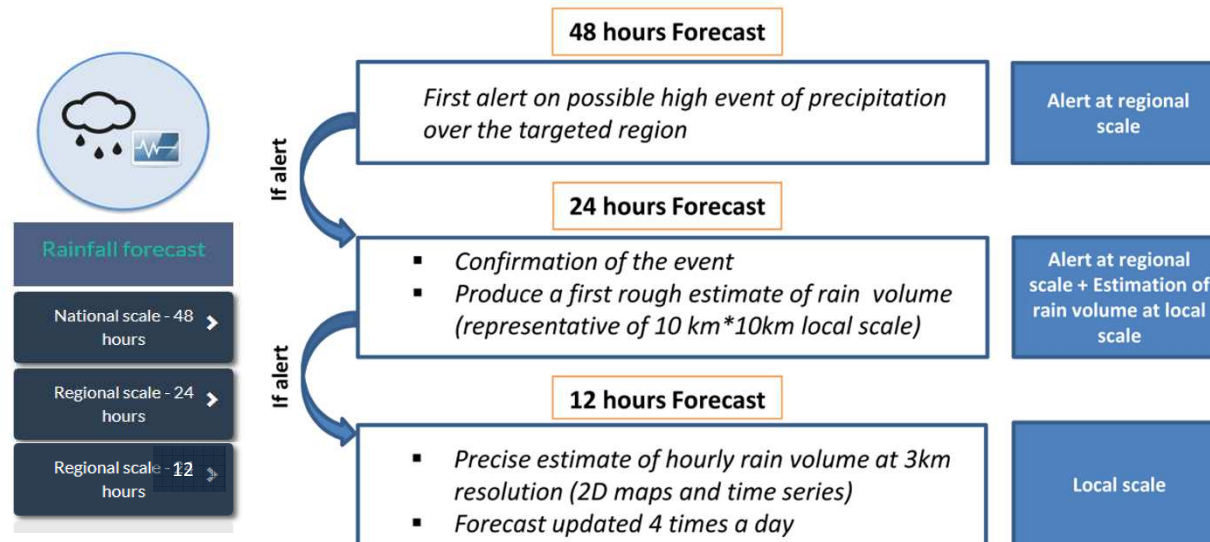
➔ Nécessite d'étendre le champ d'application et/ou les performances actuelles d'anaWEGE

Systeme d'alerte

Wat-Ener-Cast:

- Service dédié aux opérateurs de réseaux d'eaux
- Améliorer la gestion de crise en cas de fortes précipitations via la mise en place d'alerte

A gradual forecast information system



- Fourniture de cartes d'exposition via une plateforme web
- Diffusion de l'alerte par e-mail et SMS
- Possibilité d'alimenter les plateformes clients à l'aide des données produites

Module de prévision à 24h

- Module développé par CMCC
- Basé sur la méthode des analogues
 - Outil plus avancé qu' AnaWEGE sur la prévision court terme de pluie
- Fourni par CMCC en « boîte noire »
- Coût important lié à la configuration de l'outil sur chaque nouveau site
- Modèle économique non satisfaisant

Propositions ou approches en cours:

- France: 3 sites en IDF et région PACA
- Maroc: 3 sites
- Mexique
- Chili
- Brésil
- Chine

Suites envisagées

- Dans le cadre du projet GTI:
 - Tests complémentaires sur le travail déjà effectué
 - Nouvelle variable: précipitation
 - Nouvelles régions
 - Optimisation de la recherche d'analogues
 - Comparaison avec les prévisions de Météo-France
- Après le projet:
 - Projet CAFE (Marie-Curie): financement d'une thèse pour le développement d'AnaWEGE sur la thématique précipitation
 - ➔ Très bien évalué mais rejeté. Nouvelle soumission potentielle
 - Projet de nowcasting, montage d'un projet KIC climat?



MERCI



- Continuous Rank Probability Score (CRPS):

$$CRPS(x_t) = \frac{1}{N} \sum_{t=1}^N \int_y (F_t(y) - \mathbf{1}(y - x_t))^2 dy$$

$$F_t(y) = \Pr(x \leq y)$$

- Continuous Rank Probability Skill Score (CRPSS):

$$CRPSS = 1 - \frac{CRPS}{CRPS_{ref}}$$

Circulation analogues (1)

- *Reference* database **R**, containing consistent pressure (SLP and/or geopotential heights), temperature, precipitation etc. data during a reference period of observations
 - E.g. Reanalysis data for a fixed period, model control simulation
- *Target* dataset **T**, with only pressure data (SLP or geopotential height)
 - E.g. Observation during a period outside of the reference

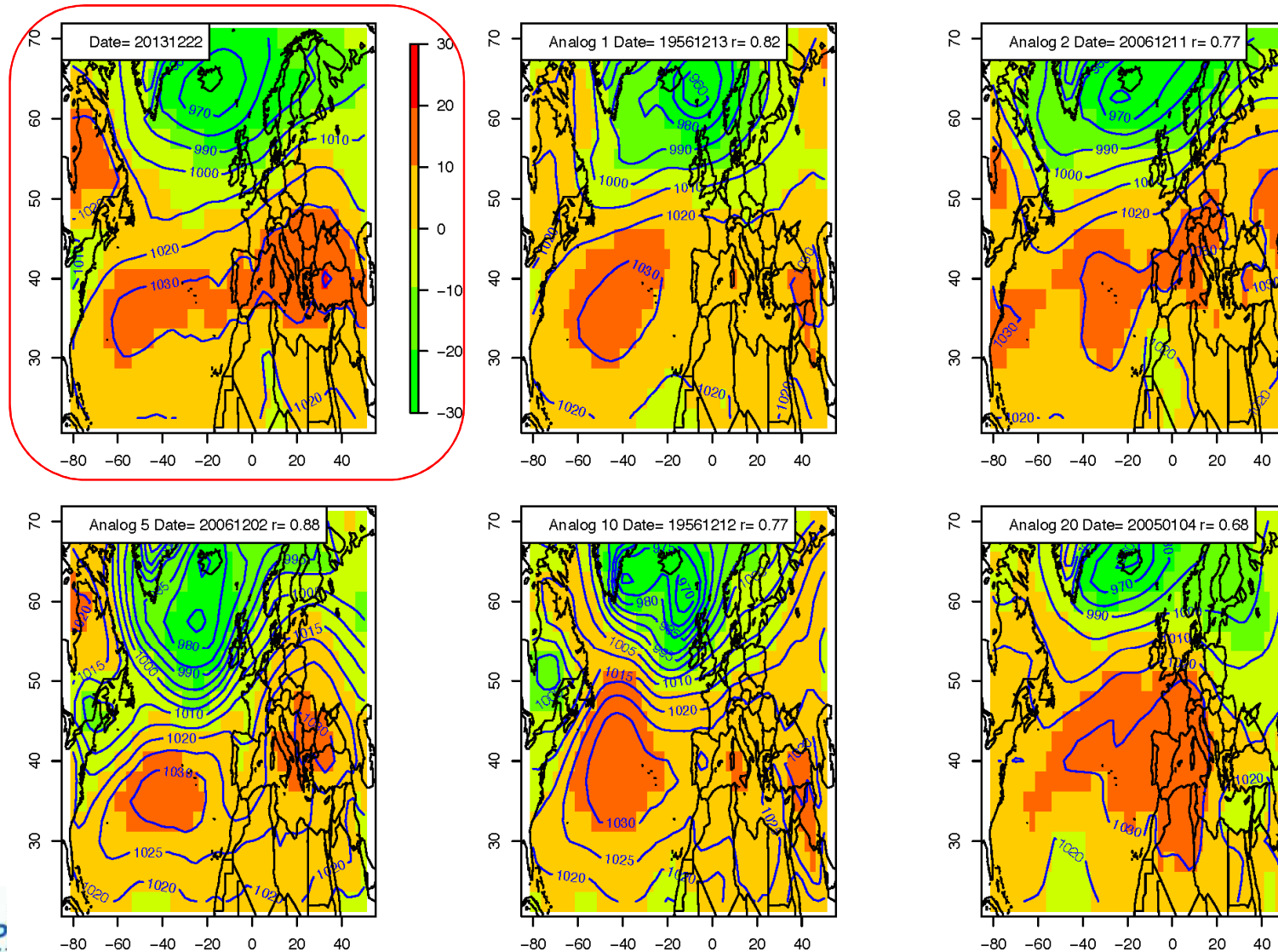
Circulation analogues (2)

- We want to infer the value of climate variables (e.g. T, Prec., Wind speed) in the dataset **T**, from information in the database **R**.
- For each day in **T**, find best analogues of pressure in **R**.
 - Minimize distance (Euclidean, Mahalanobis...)
 - Maximize spatial correlation (rank)

Circulation analogues (3)

- Use of daily sea-level pressure (SLP) from NCEP reanalyses
- For all days between Jan. 1st 1948 and March 31st 2013, pick the 10 days within 30 calendar days but different year with the closest SLP:
 - largest correlation (rank or linear)
 - Smallest Euclidean distance
- NOTE: analogues for day d , or sequences of $(d-1, d, d+1)$ to preserve local trajectories

Example: Storm Dirk



Analogue weather generators: **AnaWEGE**

- **Static**

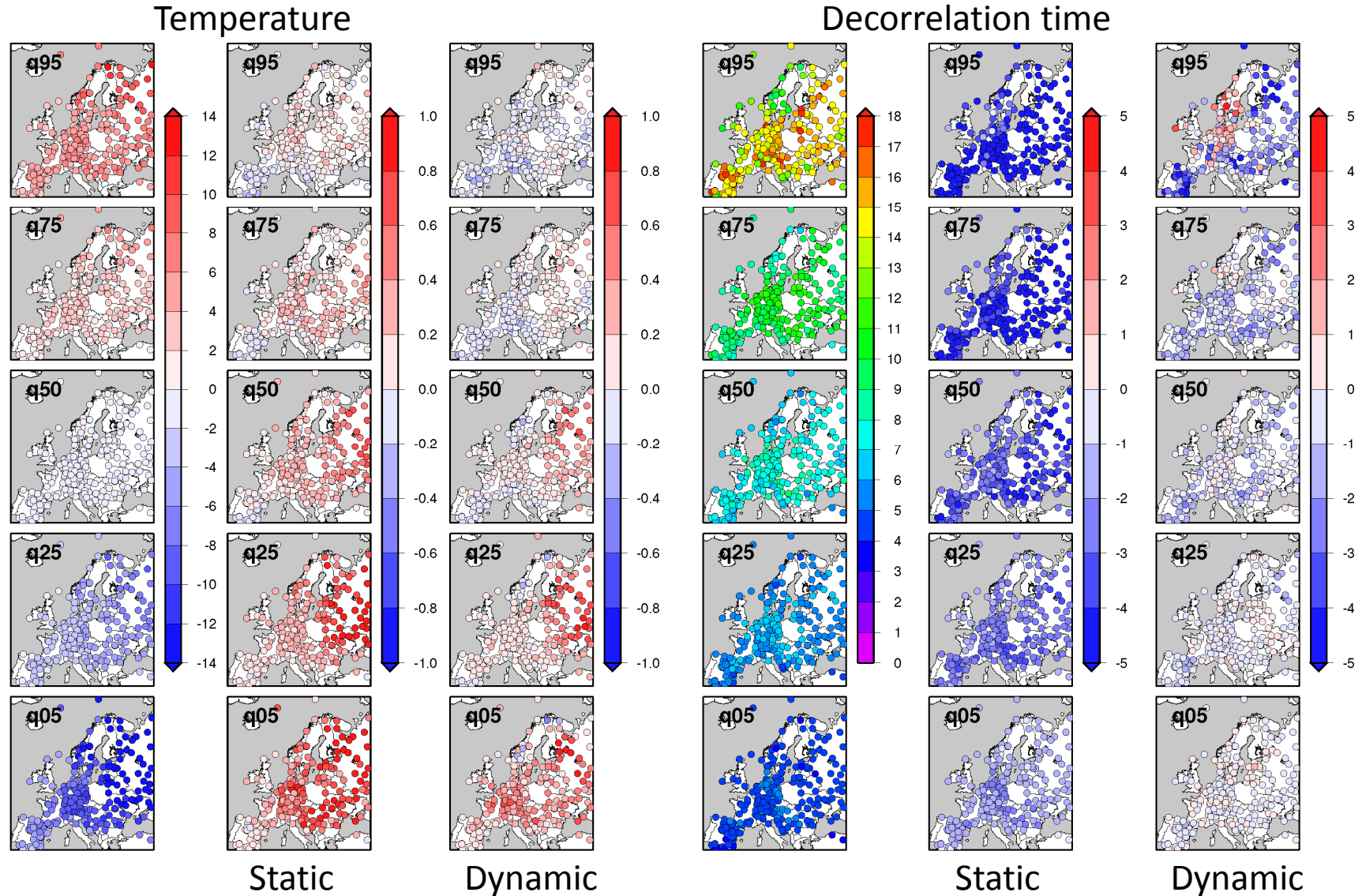
Each day ($d=yyymmdd$) is replaced by one of its 20 best analogues. The probability of drawing an analogue d' is proportional to the correlation of $SLP(d)$ and $SLP(d')$

- **Dynamic**

For each day $d (=yyymmdd)$, the next day is chosen among $yyymm(dd+1)$ and its 20 best analogues

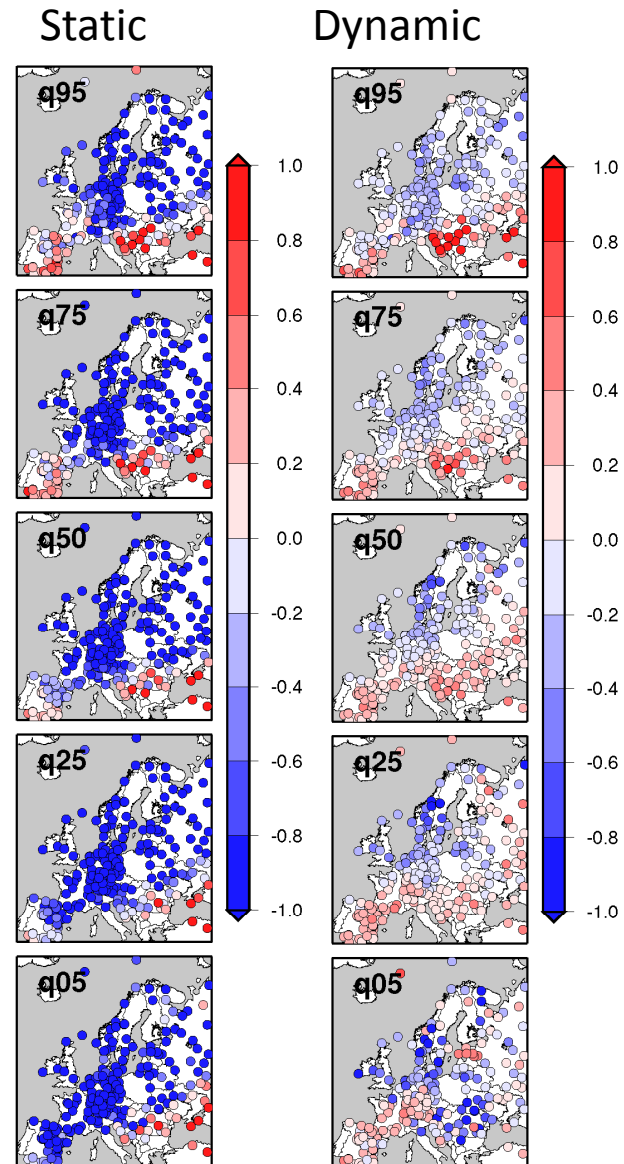
Weight of probabilities proportional to correlation and calendar distance to desired simulated date

Winter initial conditions



Cold winter initial conditions

Initial conditions on Dec. 21 2009



Features

- Interesting behavior for cold spells
- Built in spatial consistence
- Open source R code

Budget

- 37k€
- Student internships, computers, ARIA engineers

Project Goals (2)

- If interesting outcomes are identified:
 - Publications
 - PhD thesis
 - *Market study, licensing of software*