

# PINELINE

## Une base de données couplée à une application web pour le suivi en continu de la dynamique temporelle de croissance radiale du Pin maritime en Aquitaine

### Correspondance

[raphael.segura@inrae.fr](mailto:raphael.segura@inrae.fr)

[francois.ehrenmann@inrae.fr](mailto:francois.ehrenmann@inrae.fr)

Raphaël. Ségura<sup>1</sup>

Léa Antunez<sup>1</sup>

Céline Meredieu<sup>1</sup>

Thierry Labbé<sup>2</sup>

Jean-Marc Gion<sup>1,3</sup>

Laurent Bouffier<sup>1</sup>

François Ehrenmann<sup>1</sup>

### Résumé.

L'UMR BIOGECO (Campus forêt bois de Cestas Pierroton) étudie la dynamique temporelle de croissance et de statut hydrique des arbres grâce à des dispositifs instrumentés à l'aide de micro-dendromètres électroniques couplés à des capteurs de suivi environnementaux. L'ampleur de ce réseau rendait compliquée la gestion, à la fois, des dispositifs instrumentés et des données générées par l'ensemble des capteurs installés. L'ensemble des acteurs de ce réseau, chercheurs, ingénieurs et techniciens, a donc été impliqué dans l'élaboration du cahier des charges, à l'occasion de l'encadrement d'une étudiante, afin de développer une application web couplée à une base de données : PINELINE. Cet outil informatique permet de stocker et de mettre à disposition les données de phénotypage à haut débit des micro-variations radiales du tronc et les variables environnementales, pour des populations de pin maritime dans la forêt des Landes de Gascogne. Depuis 2011, au sein de l'UMR BIOGECO, plusieurs projets de recherche ont permis d'équiper un réseau de 5 dispositifs expérimentaux, pour étudier la réponse des arbres aux contraintes abiotiques, en fonction de leurs caractéristiques génétiques et des pratiques sylvicoles. Ces équipements permettent le suivi temporel automatisé des variations radiales des troncs des arbres, avec une précision au micromètre, et sont associés à des capteurs d'enregistrement de variables environnementales associées (humidité du sol, hauteur de la nappe, température, rayonnement incident), avec une fréquence de mesures horaire ou sub-horaire sur plusieurs années.

Le soutien du LabEx COTE a permis de lever les freins à la gestion technique de tels dispositifs instrumentés et le développement d'une application web associée à une base de données spécifiquement dédiée à l'enregistrement des mesures phénotypiques et environnementales à haut-débit et leur mise à disposition de la communauté scientifique.

### Mots-clés

Application web, micro-dendromètre, croissance radiale, données environnementales, statut hydrique, base de données, PostgreSQL, CakePHP.

1 BIOGECO, INRAE, Univ. Bordeaux, Cestas, France.

2 EFP, INRAE, Cestas, France.

3 UMR AGAP, CIRAD, INRAE, Univ Montpellier, France.

# PINELINE

## A database coupled with a web application to continuously monitor the temporal dynamics of the radial growth of the Maritime Pine in Aquitaine

### Correspondance

[raphael.segura@inrae.fr](mailto:raphael.segura@inrae.fr)

[francois.ehrenmann@inrae.fr](mailto:francois.ehrenmann@inrae.fr)

Raphaël. Ségura<sup>1</sup>

L. Antunez<sup>1</sup>

C. Meredieu<sup>1</sup>

T. Labbé<sup>2</sup>

JM. Gion<sup>1,3</sup>

L. Bouffier<sup>1</sup>

François Ehrenmann<sup>1</sup>

### Abstract.

The UMR BIOGECO (Forest wood campus of Cestas Pierroton) studies the temporal dynamics of the growth and hydric state of trees using systems that include instruments such as electronic micro-dendrometers coupled with environmental monitoring sensors. The size of this network complicates the management of both the arrays of instruments and the data generated by all the sensors installed. The participants in this network, researchers, engineers and technicians, have all been involved in drawing up the specifications, on the occasion of supervising a student, necessary to develop a web application coupled with a database: PINELINE. This computer tool, developed with the help of LabEx COTE, allows storing and making available the high throughput phenotyping data of radial micro-variations of trunks and environmental variables, for maritime pine populations in the forest of the Landes de Gascogne. Since 2011, within the UMR BIOGECO, several research projects have led to setting up a network of five experimental systems to study the response of trees to abiotic stresses as a function of their genetic characteristics and forestry practices. These systems ensure the automated temporal monitoring of the radial variations of tree trunks with micrometric precision and are linked to sensors that record their associated environmental variables (soil humidity, water table height, temperature, incident radiation), with hourly and sub-hourly frequencies over several years.

Support from the LabEx COTE led to the removal of barriers to the technical management of instrumented set-ups and the development of a web application associated with a database specifically dedicated to recording high throughput phenotype and environmental measurements, and making them available to the scientific community.

### Keywords

Application web, micro-dendromètre, croissance radiale, données environnementales, statut hydrique, base de données, PostgreSQL, CakePHP.

---

1 BIOGECO, INRAE, Univ. Bordeaux, Cestas, France.

2 EFP, INRAE, Cestas, France.

3 UMR AGAP, CIRAD, INRAE, Univ Montpellier, France.

## Introduction

Dans le cadre des recherches menées en génétique et en sylviculture sur le pin maritime, l'UMR BIOGECO, située sur le Campus forêt bois de Cestas Pierroton, développe l'étude de la dynamique temporelle de croissance et du statut hydrique des arbres grâce à l'utilisation de micro-dendromètres qui, déployés sous forme de réseau, permettent de suivre simultanément plusieurs centaines d'individus. Ces réseaux de capteurs couplés à l'enregistrement des variations environnementales permettent d'analyser le déterminisme génétique de l'adaptation aux contraintes abiotiques des arbres forestiers. Le premier réseau de capteurs à grande échelle a d'abord été réalisé sur 250 individus appartenant à trois familles de plein-frères issus de trois écotypes de pin maritime (landais, corse et marocain) (2011, Projet FEDER ABIOPEN). Puis, ce suivi en continu a été étendu, à partir de 2012, à 3 dispositifs du GIS « Groupe Pin maritime du Futur » et 1 dispositif du GIS « Coop de données » (projet région Fortius et projet ANR WUETree). L'objectif de ces nouveaux dispositifs instrumentés est de pouvoir tracer des normes de réactions pour différentes unités génétiques ou différents scénarii sylvicoles, afin d'intégrer des critères de sélection relatifs à l'adaptation au stress hydrique, dans le cadre du programme de création variétale. Au final, nous avons équipé, à ce jour, 450 arbres repartis sur ces 5 dispositifs. Ces derniers sont également finement caractérisés pour l'humidité du sol (plusieurs points de mesure et profondeurs au sein de chaque dispositif) et le niveau de la nappe phréatique. Par ailleurs, l'installation de stations météorologiques situées à proximité permet de disposer des variables climatiques locales : température de l'air, hygrométrie, rayonnement, précipitations, vitesse du vent. L'ensemble de ces mesures, micro-variations radiales du tronc, humidité du sol, hauteur de la nappe, variables climatiques, est réalisé à une fréquence élevée, à savoir toutes les heures ou toutes les 5 minutes selon le dispositif, rendue possible par l'automatisation de leur acquisition. Face à l'explosion des données numériques issues de l'étude de la dynamique de croissance à long terme chez les arbres, l'unité est confrontée à un manque d'outils informatiques facilitant leur stockage, leur mise à disposition et leur traitement. Par ailleurs, le développement de ce réseau de phénotypage haut débit nécessite obligatoirement la mise en place d'outils de gestion des dispositifs pour faciliter le travail de l'opérateur, car l'instrumentation de ces dispositifs nécessite un suivi technique conséquent. Le projet PINELINE soutenu par le LabEx COTE a permis de développer un outil informatique basé sur une application web couplée à une base de données pour l'étude de la dy-

namique temporelle de croissance et statut hydrique des arbres. Tous les acteurs de ce réseau qu'ils soient techniciens, ingénieurs ou scientifiques ont été impliqués dans l'encadrement d'une étudiante en master bioinformatique ainsi que dans l'établissement du cahier des charges de l'application dans le but de répondre aux objectifs suivants :

- La maintenance d'un si grand nombre de capteurs nécessite un accès rapide et simplifié aux données brutes afin de détecter d'éventuels dysfonctionnements et la mise en place d'actions correctives par le technicien en charge de la surveillance du réseau.
  - La curation et validation des données doit pouvoir se faire grâce à une interface qui permet de les traiter soit pour une période de temps définie par l'opérateur soit pour tout ou partie des capteurs d'un dispositif. Elle doit pouvoir se faire, en outre, grâce à des critères paramétrables pour permettre la résolution d'un maximum de problèmes rencontrés sur l'acquisition des mesures.
  - Le stockage des données et leur accès à la communauté scientifique doit être possible grâce à une base de données et une interface web avec différents droits associés.
- Après une présentation de l'architecture de la plateforme de phénotypage haut débit et des mesures associées, cet article décrit l'application PINELINE et sa mise en œuvre.

## Description des dispositifs expérimentaux

### Le type de dispositifs

Cinq dispositifs ont été instrumentés dans le cadre de différents projets. Le nombre d'arbres équipés de micro-dendromètres, le nombre de mesures environnementales, la fréquence des enregistrements et la durée du suivi dépend des dispositifs (Tableau 1). Les objectifs de recherche varient également selon les dispositifs.

Le dispositif « ABIOPEN » a été planté en 2008 et est constitué de 3 familles de plein-frères, d'environ 200 arbres, obtenues par croisement de 3 parents d'écotypes différents (écotype : Landes, Corse et Maroc). Ce type de dispositif permet d'étudier l'architecture génétique de caractères (analyses QTL).

Les dispositifs du GIS « Groupe Pin maritime du futur » GIS3-1 (lande sèche) et GIS3-3 (lande humide) ont été installés en 1996. Ils comprennent environ 150 familles de demi-frères (issues d'un croisement entre une mère et un mélange de 40 pollens) avec 35 individus par famille et par site. Ces dispositifs de descendance permettent d'étudier les paramètres génétiques (variabilité, héritabilité, corrélations) de caractères d'intérêt.

Le dispositif « 9533M1-Castillonville-sylviculture » a été planté en 1995. Sur plus de trois hectares, quatorze placettes permettent de comparer la croissance des arbres issus de deux niveaux d'amélioration génétique, en particulier la variété VF1, et soumis à sept scénarios contrastés de densité du peuplement dès la plantation (de 416 à 1 666 arbres par hectare à l'installation). Ce dispositif permet d'évaluer la dynamique de croissance des arbres, modulée par les intensités de récoltes partielles, dans la perspective de l'adaptation de la gestion aux changements environnementaux.

Le dispositif « Castillonville-génétique » a été installé en 1989 afin de comparer des unités génétiques en petites placettes, en particulier les premières variétés améliorées de pin maritime (VF1) par rapport au matériel non amélioré.

## Les données expérimentales

### Les micro-variations du tronc

Les micro-dendromètres utilisés pour le suivi des variations radiales du tronc utilisent le principe du potentiomètre linéaire. Un potentiomètre est un type de résistance variable à trois bornes, dont une est reliée à un curseur se déplaçant sur une piste résistante terminée par les deux autres bornes (Figure 1). Ce système permet de recueillir, entre la borne reliée au curseur et une des deux autres bornes, une tension qui dépend de la position du curseur et de la tension à laquelle est soumise la résistance. Dans notre cas, l'extrémité de la tige métallique située en contact avec le tronc de

**Tableau 1. Descriptif des différents dispositifs installés dans le cadre du réseau de phénotypage haut débit**

Nom du dispositif	Com-mune	Type de dispositifs	Micro-variation radiale mesurée toutes les 5 min	Période de suivi	Variables environnementales mesurées toutes les 5 minutes		
					Humidité du sol	Nappe phréatique	Station météorologique
ABIOGEN	33610, Cestas	3 familles de plein-frères (croisements inter-écotypes)	Landes x Corse (120 arbres) Landes x Maroc (120 arbres)	2011 - en cours	16 sondes TDR (15, 45 et 70 cm)	x	Pierroton
GIS3-1	40210, Escource	Familles de demi-frères	75 arbres (25 familles avec 3 arbres par famille)	2015 - 2020	20 sondes TDR (45 et 70 cm)	1 piézomètre	Pontenx-les-Forges
GIS3-3	33610, Cestas	Familles de demi-frères	75 arbres (25 familles avec 3 arbres par famille)	2015 - 2020	20 sondes TDR (45 et 70 cm)	1 piézomètre	Cestas
9533M1-Castillonville Sylviculture	33610, Cestas	Comparaison de différents scénarios de densité pour un niveau d'amélioration	20 arbres dans 4 placettes (3 ou 4 arbres équipés par placette)	2015 - en cours	42 sondes TDR (25 à 70cm) réparties dans 6 placettes	1 piézomètre	Cestas - Castillonville
Castillonville Génétique	33610, Cestas	Comparaison de variétés non améliorées et améliorées	20 arbres	2015 - en cours	42 sondes TDR (25 à 70cm)	1 piézomètre	Cestas - Castillonville

$$V_{\text{mesure}} = \frac{R_a}{R_a + R_b} \times V_{\text{ref}}$$

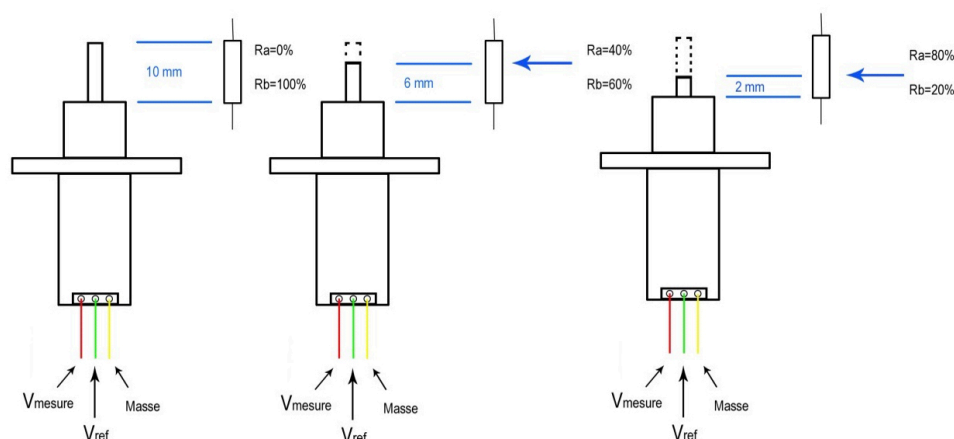


Figure 1. Principe de fonctionnement du capteur de positionnement type potentiomètre linéaire

l'arbre fera varier la valeur de la résistance. Dans le cas des micro-dendromètres, une tension de référence régulée  $V_{\text{ref}}$  est appliquée au capteur et la tension mesurée  $V_{\text{mesure}}$  résulte du rapport entre les résistances variables.

Afin d'être capable de traduire les variations de tension en déplacement physique de la pointe du dendromètre, une calibration est réalisée le long de la gamme de déplacement [0-10mm]. Cette étape permet d'obtenir une droite d'ajustement de type  $y = a x + b$  pour chaque micro-dendromètre et permet l'inter-comparaison des différents capteurs installés sur les troncs.

### Température et statut hydrique du sol

Chaque dispositif est équipé de sondes réflectométriques de teneur en eau du sol (CS650, Campbell Scientific Ltd), installées à différentes profondeurs (25-45-70 cm). Ces sondes sont disposées dans des fosses à proximité des arbres suivis pour leurs micro-variations radiales. Les sondes mesurent le temps de propagation et l'atténuation du signal électrique ainsi que la température dans un volume de sol situé autour des électrodes métalliques. La mesure de l'atténuation du signal permet de mesurer la conductivité électrique et ainsi de corriger le temps pour lequel la réflexion de l'onde est détectée. Cette correction permet d'améliorer la mesure du temps de propagation. La vitesse de propagation de l'onde est modulée par la permittivité du matériau traversé, principalement influencée par la présence d'eau autour des électrodes. Le processeur intégré à la sonde permet ensuite d'estimer la teneur en eau volumétrique en utilisant la formule de Topp développée en 1980.

En outre, un piézomètre a été installé sur chaque dispositif (à l'exception de « ABIOMEN ») pour le suivi de la hauteur de

la nappe phréatique perchée, élément primordial dans la compréhension de l'alimentation hydrique des arbres dans le Massif des Landes de Gascogne. Les profondeurs de forage ont été adaptées à la topographie des lieux allant de 4 à 15 m. Des capteurs de pression submersibles (CS450, Campbell Scientific Ltd) sont insérés dans des crépines et tubes PVC de 80 à 120 mm de diamètre et enregistrent les variations de la hauteur de la nappe.

### Informations météorologiques

Sur chaque dispositif, les mesures de la température et l'humidité de l'air sont réalisées à l'aide de capteurs CS215, Campbell Scientific Ltd. Le dispositif comprend un élément de détection capacitif en polymère pour l'humidité relative et un capteur de température à semi-conducteur.

Deux centrales météorologiques ont été installées à proximité géographique des dispositifs instrumentés. Elles enregistrent les variables environnementales suivantes :

- Température et humidité de l'air (CS215, Campbell Scientific Ltd)
- Rayonnement photosynthétique (SKP 215, Skye instruments)
- Vitesse du vent (A100LK, Vector instruments)
- Pluviométrie (ARG100, pluviomètre à auget basculant, Campbell Scientific Ltd).

### Enregistrement et transfert des données

L'instrumentation de chacun des 5 dispositifs repose sur la même architecture. Le dispositif est découpé en blocs (entre 3 et 6) chacun équipé par une centrale de mesures CR1000 et/ou un multiplexeur AM16/32 (Campbell Scien-

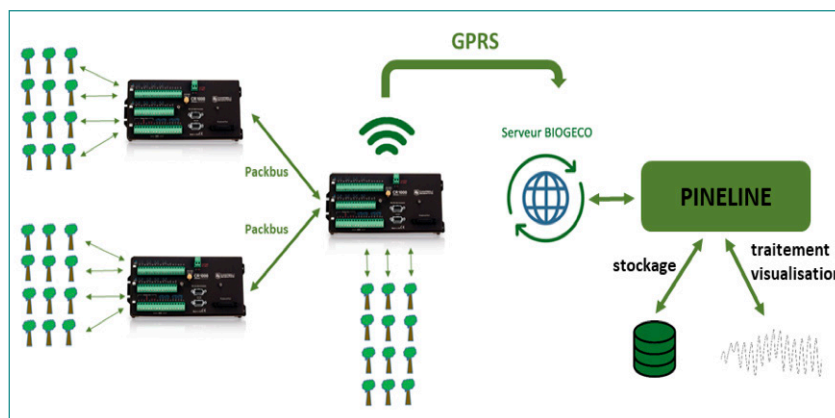


Figure 2. Schéma simplifié du dataflow PINELINE

tific Ltd). Les différents capteurs sont scannés à intervalle de temps régulier allant de 5 à 60 minutes selon les cas. La communication se fait via un réseau Packbus (Campbell Scientific Ltd) reliant centrales de mesure et multiplexeurs, soit à l'aide d'interfaces filaires RS485 soit liaison radio 2.4GHz (Figure 2).

Quotidiennement et grâce au protocole GPRS, les centrales maîtres équipées de modem de transmission permettent la collecte à distance des données sur un serveur local équipé du logiciel Loggernet Admin (Campbell Scientific Ltd). Les fichiers de données peuvent alors être insérés dans la base de données via l'application web PINELINE.

## PINELINE : une application web couplée à une base de données

### La Base de données

Les développements de la base de données PINELINE et son interface web ont été entamés lors du stage de Master 2 Bioinformatique réalisé par Léa Antunes au sein de l'équipe EGF (Écologie et Génomique Fonctionnelles) de BIOGECO, dans le cadre d'un financement du Laboratoire d'Excellence COTE (LabEx COTE). La conception de la plateforme était prévue sur une durée de trois ans et se composait de quatre étapes principales :

- création de la base de données,
- création d'un site web intégré à un portail de bases de données,
- mise en place d'une interface « opérateur » pour le suivi et l'intégration des données en « temps continu »,
- mise à disposition d'une interface « utilisateur » pour rendre la plateforme accessible à la communauté scientifique.

PINELINE a donc pour finalité d'être une application web ouverte à toute la communauté scientifique et de disposer de liens étroits avec des portails de bases de données existants

tels que Pinus Portal. L'interface de l'application se veut la plus intuitive possible pour, à la fois, faciliter l'interrogation et l'accès aux données et aider à son alimentation. Pour cette fonctionnalité d'alimentation de la BDD, il est possible pour l'opérateur d'ajouter les dernières données de mesure concernant un dispositif, ou encore de créer un nouveau dispositif et y ajouter ses caractéristiques descriptives.

### Du modèle conceptuel de données au schéma relationnel

Un Modèle Conceptuel de Données (MCD), ou modèle entité-association, est une représentation de la structure du système d'information, prenant en compte les relations entre les différentes entités. Une entité est définie par un ensemble de propriétés, et chaque combinaison de propriétés doit être unique. Pour cela, une propriété particulière, appelée clé, permet de l'identifier. Une association est un lien sémantique entre une ou plusieurs entités ; cette relation est également caractérisée par sa cardinalité qui correspond au nombre d'occurrences d'une entité dans une association. Il est possible de distinguer quatre types de données descriptives dans le cas d'un plateau de suivi équipé d'une centrale d'acquisition : le matériel végétal, la station météorologique, les capteurs sol et température/RH ainsi que et les capteurs des micro-variations radiales (Figure 3).

Chaque entité d'un dispositif de suivi de croissance est décomposée pour réaliser la base du MCD. Il a été convenu de représenter un dispositif expérimental dans sa globalité, c'est-à-dire incluant tous les arbres plantés et non exclusivement ceux qui sont suivis avec des microdendromètres, ceci afin de permettre une évolution de la plateforme en cas d'acquisition de nouvelles données de phénotypage. Le MCD réalisé a permis, par la suite, de construire le schéma relationnel de la base de données (Figure 4). Ainsi, la réalisation de ce schéma relationnel global permet l'orga-

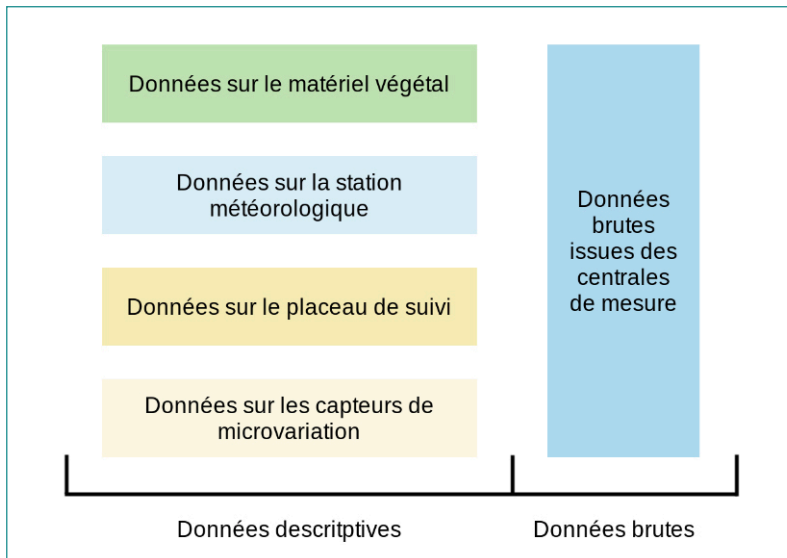


Figure 3. Représentation schématique des différentes données descriptives et brutes pour la création du MCD du projet PINELINE

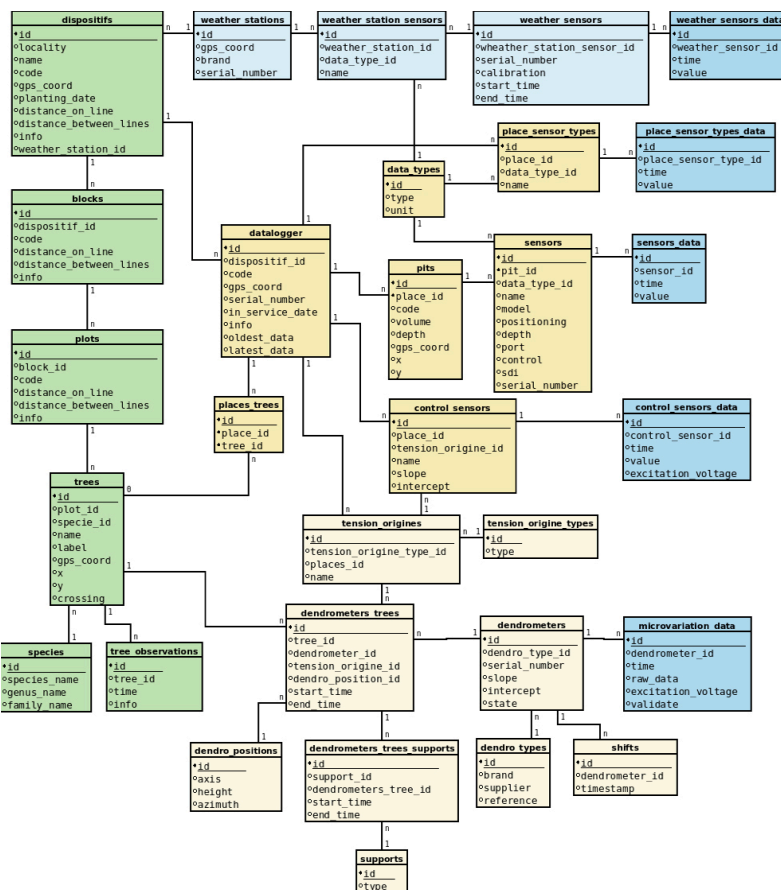


Figure 4. Schéma relationnel de la base de données PINELINE

nisation et la structuration des entités et des informations recueillies par le personnel en charge du projet. Ce schéma est à la base de la plateforme PINELINE et permet de stocker et d'organiser les données pour faciliter les recherches, mais également d'archiver de manière pérenne les données pour le travail de l'équipe.

### Stratégie pour l'insertion des données (Figure 5)

Afin d'alimenter la base de données, il est nécessaire que les données descriptives soient au préalable renseignées pour que chaque donnée brute soit associée à la bonne entité. Les données brutes provenant des centrales sont fournies sous un format tableur CSV, ce qui en facilite le « parsing », via des bibliothèques ou modules adéquats. Cependant, chaque centrale d'enregistrement produit un fichier différent, que ce soit sur le nombre de colonnes ou leur contenu. Tout d'abord, le nombre d'arbres suivis, de sondes ou encore de témoins diffère selon le plateau de suivi. De plus, l'ordre et l'identification des colonnes ne sont jamais identiques entre les cinq dispositifs existants. Différentes stratégies ont donc été pensées et testées pour l'insertion des données brutes. Finalement, la stratégie retenue consiste en la création d'un fichier de configuration par plateau

de suivi, et « écrit » par l'opérateur de la plateforme. Après avoir choisi les dispositifs et plateaux de suivi auxquels les données vont être ajoutées, une étape de description du fichier de données brutes est nécessaire. Ainsi, pour chaque colonne, l'opérateur choisit le type de données contenu dans la colonne, puis l'entité correspondante dans la base de données. Cette opération permet d'écrire un fichier de configuration contenant le numéro de la colonne, le type de données, et enfin l'identifiant de l'entité dans la base de données. Cette étape n'est effectuée qu'une seule et unique fois par l'opérateur, avant la première insertion de données brutes. Une fois créé pour un plateau de suivi, le fichier de configuration est réutilisé pour tout ajout de données pour ce plateau (la structure des fichiers de données brutes ne variant pas au cours du temps). Cette stratégie, permettant un minimum de paramétrage par dispositif et plateau, a été retenue pour l'insertion des données brutes.

### Choix de conception

La conception de la plateforme PINELINE doit utiliser un système permettant son évolution et donc disposer d'un bon niveau de modularité. Le système réparti est un système dont les ressources ne sont pas centralisées afin de

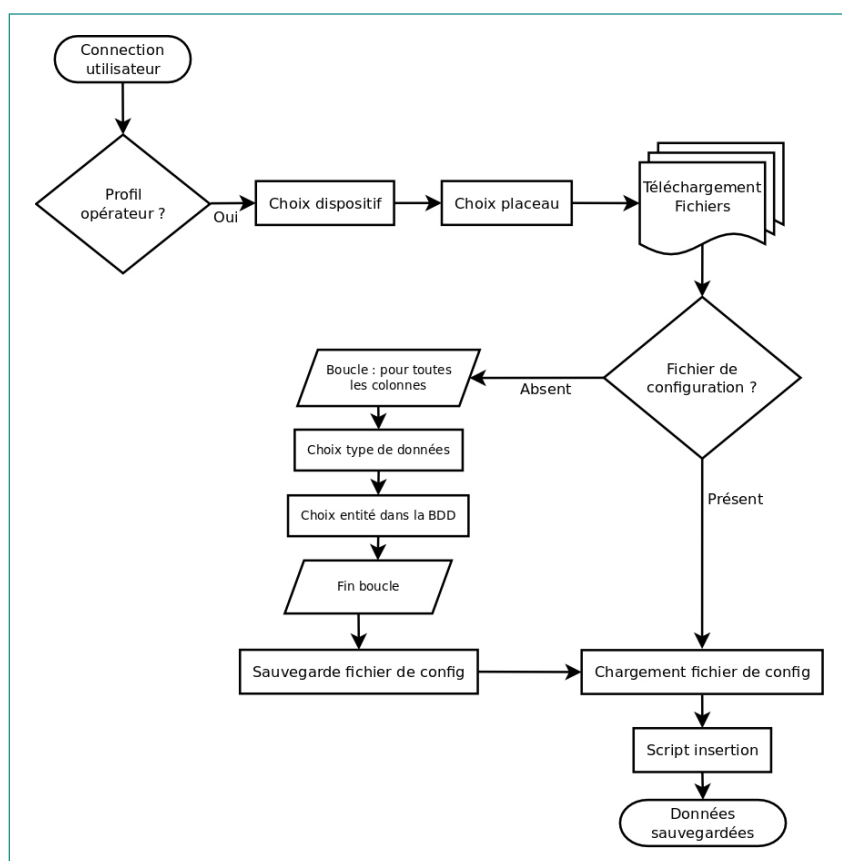


Figure 5. Algorithme de flux pour l'insertion des données brutes dans la base de données à partir de fichiers tabulés.



permettre aux utilisateurs de manipuler leurs données sans contrainte sur la localisation des éléments, et de diviser les problématiques. Ce système client-serveur comprend un client (navigateur web), un serveur applicatif (présentation, logique de l'application et gestion des ressources) et une base de données. Une unité de calcul est mise en place, pour répondre à la problématique de la volumétrie des données générées par les centrales. Celle-ci dépendra, en effet, de la fréquence de mise à jour de la base de données par les opérateurs et, surtout, du pas de temps choisi pour les mesures. De plus, les traitements sur les données brutes nécessitent un accès à cette unité de calcul.

Dans le cadre de l'application PINELINE, l'architecture trois-tiers est utilisée pour séparer la réalisation des trois parties du serveur applicatif. Tout comme dans le client-serveur, cette séparation signifie qu'il est possible de déployer chaque partie sur un serveur indépendant. Ce type d'architecture permet, dans tous les cas, une plus grande évolutivité du système. Une architecture en trois-tiers (Figure 6) se compose de trois éléments :

- un Système de Gestion de Bases de Données pour le stockage des données (SGBD PostgreSQL version 9.6),
- un serveur applicatif pour la logique applicative (Serveur web : Apache / Applicatif : Framework CakePHP 3),
- un navigateur web pour la présentation (client).

L'avantage principal d'une architecture trois-tiers est son déploiement. En effet, l'application n'est déployée que sur la partie serveur (serveur applicatif et serveur de base de données) ; le client ne nécessite qu'une installation et une

configuration minimale (installer un navigateur web compatible avec l'application). Cette facilité de déploiement est importante, car elle permet une évolution régulière du système en mettant à jour seulement le serveur applicatif. Un autre avantage de cette architecture est la sécurité : en effet, dans un système client-serveur, les clients accèdent directement à la base de données, ce qui n'est pas le cas dans une architecture trois-tiers où l'accès est effectué par le serveur applicatif. Ainsi la sécurité de l'application se fait au niveau du serveur applicatif, en gérant la liste des utilisateurs, leur mot de passe et leurs droits d'accès aux fonctionnalités de l'application.

## Description de l'application web PINELINE

### Droits utilisateurs

Il existe 4 droits utilisateurs permettant d'accéder à l'application PINELINE :

- **Administrateur** : utilisateur ayant tous les droits et accès à l'ensemble des fonctionnalités. Accès à l'ensemble des tables de la base de données en lecture/écriture. Ajout/modification/suppression d'utilisateurs.
- **Opérateur** : utilisateur ayant accès à l'ensemble des fonctionnalités, sauf les accès à l'ensemble des tables du modèle en lecture/écriture et à l'ajout/modification/suppression d'utilisateurs.
- **Utilisateur** : utilisateur ayant accès aux données de descriptions des dispositifs et des données brutes ou curées par l'intermédiaire d'un formulaire de recherche. Possibilité de télécharger les données mises à disposition au public.
- **Visiteur** : utilisateur ayant accès aux données de descriptions des dispositifs et des données de curées par l'intermédiaire d'un formulaire de recherche. Pas de téléchargement possible.

### Page d'accueil

L'accès aux différents menus de l'application requiert obligatoirement une authentification. La page d'accueil contient, dans sa barre supérieure, un formulaire pour permettre à l'utilisateur de s'identifier, en choisissant au préalable l'espèce pour laquelle il veut consulter les informations disponibles en base. Pineline a, en effet, vocation à intégrer des données de dispositifs pour différentes espèces (au 1<sup>er</sup> Août 2021, la base de données contient des dispositifs de pin maritime et d'eucalyptus). Sur la page d'accueil, on retrouve un résumé rapide des données présentes dans la base.

Trois sections principales sont disponibles la page d'accueil (Figure 7) :

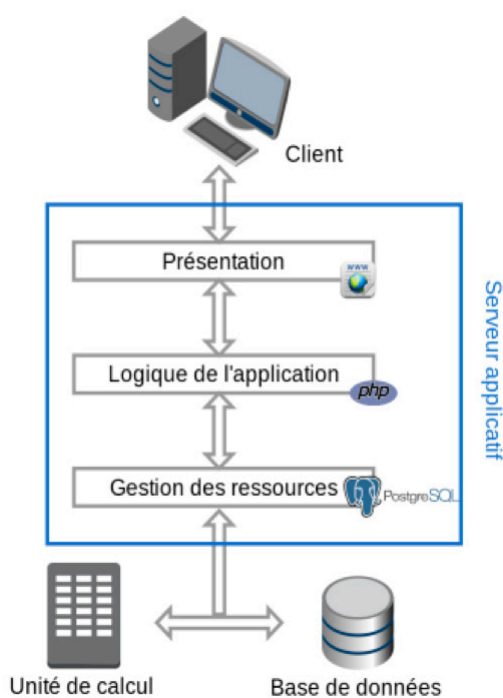


Figure 6. Schéma du système réparti utilisé pour l'application PINELINE

- **Contenu de la base** : section fournissant des métadonnées sur les dispositifs contenus dans la base (nom des dispositifs, description et diverses métriques) (Tableau 1).

- **Recherche** : section permettant l'accès aux données concernant le statut hydrique du sol et météorologique, d'une part, et, d'autre part, au script de traitement des micro-variations radiales pour une période donnée.

- **Opérateur** : section permettant à l'opérateur, en charge du suivi des dispositifs, d'administrer et d'alimenter la base de données PINELINE et notamment :

- \* créer de nouveaux dispositifs ou modifier les dispositifs existants par l'ajout en base de données descriptives ;

- \* charger des données liées aux dispositifs enregistrés dans PINELINE ;

- \* enregistrer différentes opérations de maintenance, comme le remplacement de micro-dendromètres.

Au-delà du fait de pouvoir gérer, organiser et stocker l'ensemble des données relatives à l'organisation des dispositifs et celles générées par les micro-dendromètres et autres capteurs, l'application web/base de données permet aussi aux opérateurs terrain de pouvoir simplifier considérablement leurs tâches relatives à la gestion des dispositifs et des matériels de mesure associés, mais également de réaliser un suivi dynamique des données intégrées en base (Figure 8). La section opérateur permet, en effet, de relever et d'enregistrer toutes les opérations qu'il effectue, notamment les remplacements de capteur. Ces informations sont indis-

pensables à enregistrer, car chaque capteur possède une calibration qui lui est propre et qui influe par conséquent la donnée brute.

## Menu et fonctionnalités de l'appliquatif PINELINE

### Chargement des données de micro-variation radiale et environnementales

Une fois l'ensemble des données descriptives d'un dispositif renseigné, il est possible d'ajouter les données brutes (micro-variations, sol et météo) issues des fichiers collectés à distance. Ce chargement, effectué par un opérateur, se déroule de la sorte :

- **Chargement des données** : l'opérateur choisit le dispositif et le plateau de suivi (centrale de mesure) auquel il veut ajouter les données. Les fichiers sont stockés dans un répertoire de travail temporaire spécifique au dispositif choisi. Un fichier unique de données peut être inséré dans la base, mais l'insertion d'un lot de fichiers est également possible grâce à un script en langage python. Celui-ci concatène alors tous les fichiers en un seul. La structure des fichiers choisis doit néanmoins être organisée de la même façon pour les fichiers d'un même lot.

- **Création du fichier de configuration** : la stratégie retenue consiste, lors de la première insertion dans un dispositif, en la création d'un fichier de configuration par plateau

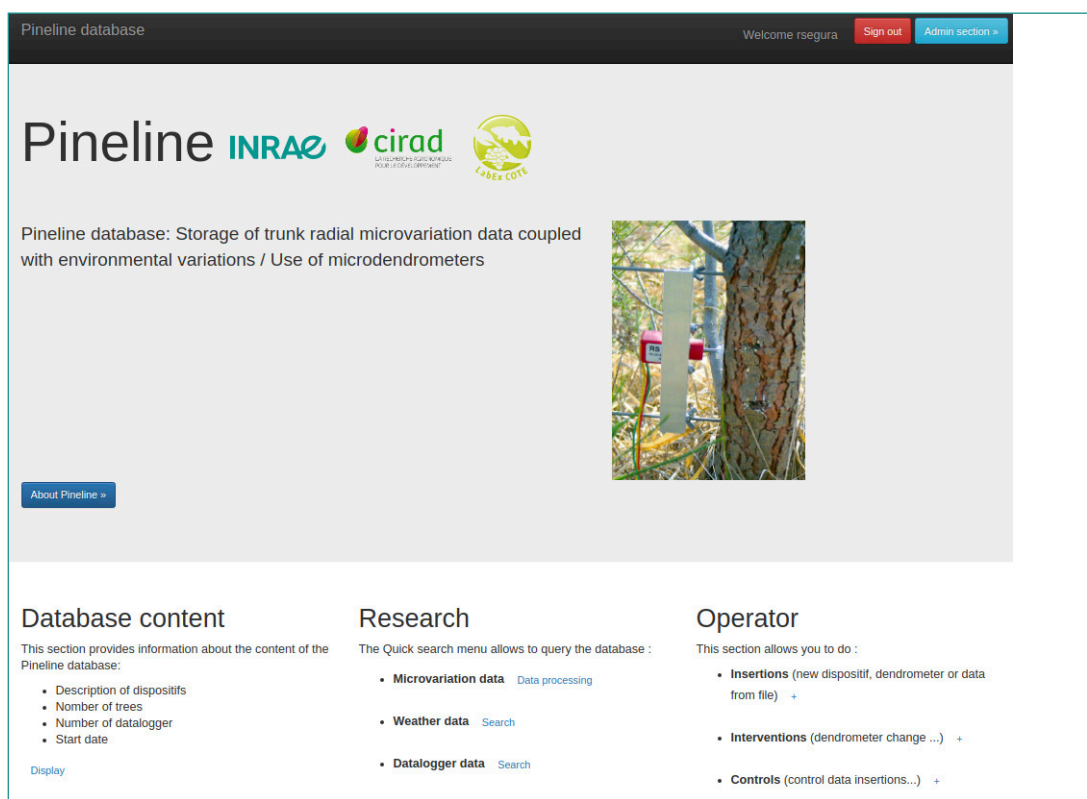


Figure 7. Page d'accueil de l'application PINELINE

Figure 8. Page d'accueil pour la saisie des métadonnées et l'insertion des données brutes

de suivi, et « écrit » par l'opérateur de la plateforme. Après avoir choisi les dispositifs et placeaux de suivi auxquels les données vont être ajoutées, une étape de description du fichier de données brutes est nécessaire. Ainsi, pour chaque colonne, l'opérateur choisit le type de données contenu dans la colonne, puis l'entité correspondante dans la base de données. Cette opération permet d'écrire un fichier de configuration contenant le numéro de la colonne, le type de données et, enfin, l'identifiant de l'entité dans la base de données. Une fois créé pour un placeau de suivi, le fichier de configuration est réutilisé pour tout ajout de données pour ce placeau (la structure des fichiers de données brutes ne variant pas au cours du temps). Cette stratégie, permettant une économie de paramétrage par dispositif et placeau, a été retenue pour l'insertion des données brutes.

### Contrôle des données de micro-variation radiale, de statut hydrique du sol et météorologiques

L'opérateur peut, à tout moment, contrôler l'ensemble des données intégrées en base, grâce à un menu dédié. Nous allons uniquement décrire ici le formulaire de recherche pour le contrôle des données de micro-variation (Figure 9). Le formulaire de contrôle permet d'accéder aux données d'un dispositif, un placeau de suivi et même un arbre unique. L'interrogation se fait grâce à des menus dérou-

lants, y compris pour les dates de début et de fin d'acquisition. Le contrôle est réalisé par l'appel en tâche de fond d'un script en langage R, script qui va permettre de charger le fichier de résultat créé par la requête SQL exécutée, et dans un second temps de créer un graphique au format PDF (Figure 10). Lors de ce processus, des alertes peuvent être affichées sur les graphiques générés (en rouge sur la figure 10) en fonction de paramètres choisis par l'opérateur. Il peut alors télécharger le fichier PDF contenant les graphiques de données brutes sélectionnées ainsi que le fichier CSV pour accéder aux données. Un fichier spécifique d'alerte permet également d'identifier les micro-dendromètres arrivés en bout de course et nécessitant d'être recalés.

### Nettoyage et traitement des données de micro-variation radiale

Grâce au menu recherche, les utilisateurs autorisés peuvent également lancer un script de nettoyage et de traitement de données de micro-variation radiale sur une période souhaitée (Figure 11). L'utilisateur pourra alors télécharger soit un fichier graphique au format PDF, soit les données sous forme de tableur CSV. Le traitement des données peut se faire pour tous les arbres d'un dispositif mais également se limiter aux arbres d'un placeau de ce dispositif, voire à un seul arbre. La finalité du script de nettoyage et de traite-

ment des données de micro-variation radiale est de traduire le signal électrique (millivolts) en un déplacement exprimé en millimètre des fluctuations du diamètre de tronc (Figure 12). Pour cela, le script utilise les informations contenues dans la table « dendrometer\_tree » pour associer la droite d'ajustement propre à chaque micro-dendromètre sur la période d'association avec l'arbre étudié. La fenêtre permet de paramétrer quelques éléments du script de nettoyage :

- Valeur seuil pour les données aberrantes qui seront supprimées.

- Valeur seuil entre 2 données successives, qui sera considérée comme normale mais signifiant que le capteur est arrivé en bout de course en raison de la croissance de l'arbre, et entraînera une correction automatique.
- Valeur seuil entre 2 données successives qui sera considérée comme inacceptable et supprimée automatiquement, comme par exemple lors de l'apparition de bruit sur le signal.
- Le coefficient de correction des données en fonction de la dilatation thermique selon le type de support du dendromètre.

The image shows two side-by-side screenshots of a web application interface titled 'Pipeline database'.

The left screenshot is titled 'Search microvariation data'. It contains several search filters:
 

- Select date start:** 2019, March, 1, 0, 00
- Select date end:** 2019, October, 1, 0, 00
- Select dispositif:** GIS3-3
- Select datalogger:** Cestas\_CES22\_Multi
- Select tree:** 3971272Dendro\_016
- Delay (minutes):** 60
- Options:**
  - Alert file (Create if a microvariation data is greater than 8000)
  - Color by:**
    - Validated data
    - Delta > 100 (between 2 values)
    - No color

 A 'Submit' button is at the bottom.

The right screenshot is titled 'Result'. It displays:
 

- Control graph:** GIS3-3
- Dispositif:** GIS3-3
- Datalogger:** Cestas\_CES22\_Multi
- Datalogger ID:** 3971272
- Download PDF:** A PDF icon with a red 'PDF' label.
- Download results:** A CSV icon with a purple 'CSV' label.
- Validated data file:** A CSV icon with a purple 'CSV' label.

Figure 9. Production des fichiers pdf et CSV pour la visualisation et le contrôles des données de micro-variations

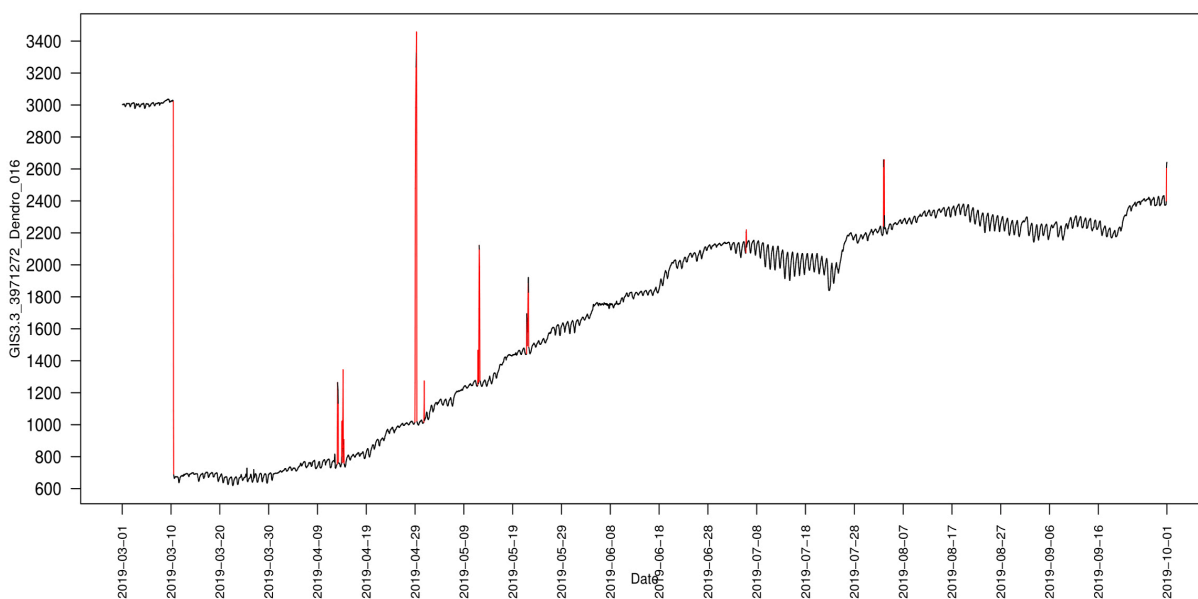


Figure 10. Exemple d'un fichier de contrôle au format pdf du micro-dendromètre 016. On observe en rouge à la fois un recalage au mois de mars ainsi que des valeurs aberrantes ponctuelles

## Interventions sur les dispositifs

Cette section permet à l'opérateur de consigner toutes les opérations effectuées sur chacun des dispositifs, notamment les changements de sondes environnementales ou de dendromètres. En effet, la calibration des micro-dendromètres est individuelle et ses coefficients doivent être renseignés par l'opérateur et rentreront en jeu dans le processus de traitement du signal. Plusieurs types d'action peuvent

être renseignés par un formulaire indiquant la date et l'heure de l'intervention et l'équipement concerné :

### Pour les dendromètres :

- le changement du support
- le recalage (nouvelle calibration)
- la modification d'un dendromètre
  - stop & change : arrêt du dendromètre et remplacement
  - stop : arrêt du dendromètre sans remplacement

Le changement d'une sonde météorologique.

Pineline database Welcome rsegura [Sign out](#) [Ad](#)

### Microvariation data processing

Select date start : 2019 March 1 0 00

Select date end : 2019 October 1 0 00

Select dispositif : GIS3-3

Select dataloger : Cestas\_CES22\_Multi

Select tree : 3971272Dendro\_016

Delay (minutes) : 60

Detect extrem aberrant value : err 10000

Detect shift correction : corr 1000

Detect aberrant value : skip 100

T° expansion correction of the supports : coeff 0.00036

Delete corr  
If shift correction detected, two values are deleted

Add sensor data in result :

air sensors

soil sensors

Figure 11. Menu de traitement des données de micro-variation

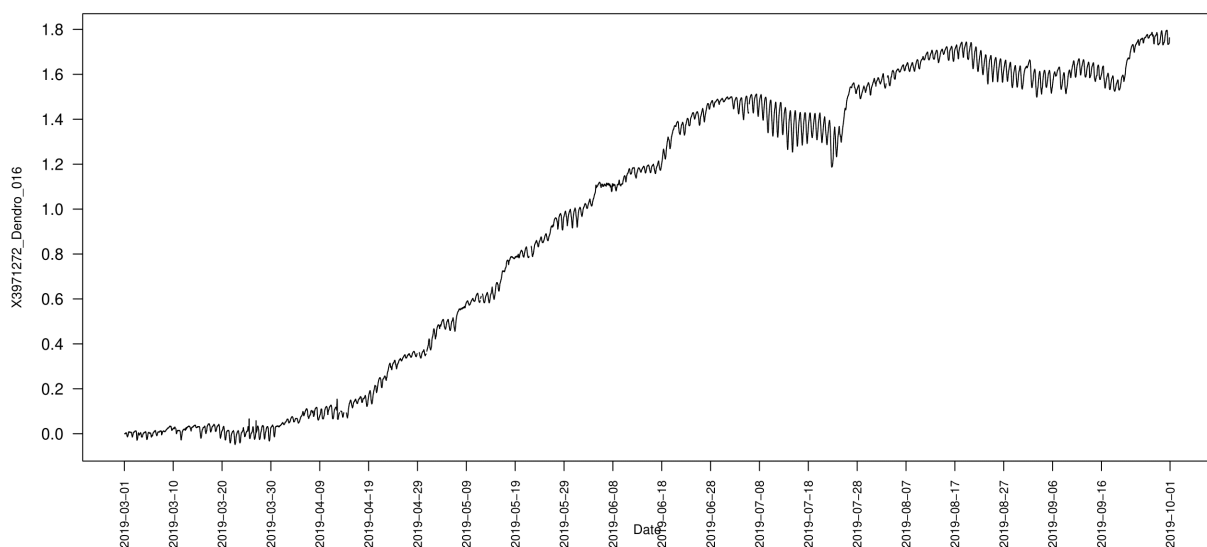


Figure 12. Exemple du signal du micro-dendromètre 016 après nettoyage et traitement automatique des données

## Conclusion

L'objectif principal du projet PINELINE, correspondant au développement d'une application web associée à une base de données spécifiquement dédiée à l'enregistrement des mesures phénotypiques et environnementales à haut-débit, a été rempli. Ce développement a permis de formaliser une approche innovante facilitant l'étude de la dynamique de croissance et du statut hydrique des arbres, couplée aux variations environnementales sur 450 arbres répartis dans 5 dispositifs. La mise en place de ce réseau de phénotypage haut débit engendre un suivi important par l'opérateur principal de ces dispositifs, suivi rendu possible grâce aux différentes fonctionnalités de l'application.

Cette application web fournit de nombreuses fonctionnalités facilitant l'intégration et le traitement des données acquises, dont notamment l'alimentation simple et rapide de la base de données par les opérateurs terrain via une intégration régulière des données brutes. Une interface de contrôle et de visualisation donne accès aux mesures brutes et permet un suivi efficace des dispositifs grâce au déclenchement rapide des interventions de maintenance adaptées aux problèmes détectés. Les dispositifs, même géographiquement éloignés de l'unité, bénéficient de cette surveillance grâce à la récolte automatique des données à distance par GPRS. La possibilité de créer des fichiers d'alerte ou de paramétrer une visualisation facilitée (en rouge selon des critères de seuil entre valeurs successives) améliore l'efficacité de l'opérateur dans la détection des problèmes. La curation des données apporte une réelle plus-value face au volume important de données générées et pouvant présenter des données aberrantes. En effet, l'opérateur détecte automatiquement, grâce aux différents paramètres de seuil, les erreurs qui peuvent se produire et les corrige automatiquement dans la plupart des cas pour tout un

lot de données. La possibilité offerte par l'application de régler ces paramètres pour un arbre individuellement permet d'ajuster si nécessaire les seuils pour ces cas particuliers à traiter.

Il reste néanmoins quelques limites comme, par exemple, le besoin de vérifier les signaux curés obtenus dans les cas particuliers où l'automatisation n'est pas efficace. De plus, la taille relativement importante de la base peut également amener une réflexion sur les outils utilisés pour stocker et interroger efficacement les données présentes en base (Nosql vs Sql).

Le développement en cours d'un menu de validation des données permettra, à terme, de ne préserver en base de production que des données curées et validées par l'équipe scientifique qui gère les données du réseau de suivi. Les données brutes en sortie de centrales seront, quant à elles, stockées et archivées dans un espace dédié, indépendamment de la base de données. De ce fait, le stockage, la manipulation, l'accès et le partage des informations générées par le réseau pourra se faire à destination de la communauté scientifique nationale et internationale.

La politique de gestion des données de l'UMR Biogeco repose, depuis fin 2021, sur le chantier « gouvernance des données » porté par l'INRAE et, également, sur le Plan de Gestion de Données d'unité. Cela concerne notamment la production de données FAIR (F = faciliter la découverte des données, stocker durablement les données et les métadonnées ; A = faciliter leur accès et/ou leur téléchargement ; I = utilisables, intelligibles ; R = réutilisables pour de futures recherches ou d'autres finalités). Ces principes FAIR sont respectés autant que possible pour ces données recueillies à haut débit, *in natura*, non reproductibles et avec une haute valeur patrimoniale. ■

# Pour aller plus loin...

Bartholomé J, Mabiála A, Burlett R, Bert D, Leplé JC, Plomion C, Gion JM (2020). The pulse of the tree is under genetic control: eucalyptus as a case study. *The Plant Journal* 103, 338–356.

Bouffier L., Segura R., Ehrenmann F., Gion J.-M. Genetic parameters of maritime pine radial growth during a specific drought event. Presented at : Forest Genetics 2017 - Forest Health and Productivity in Changing Environments, Edmonton, Canada (2017-06-26), <https://hal.inrae.fr/hal-03341365>.

Bouriaud O, Leban J-M, BertD, Deleuze C (2005) Intra-annual variations in climate influence growth and wood density of Norway spruce. *Tree Physiol* 25:651–660.

Cocozza C, Tognetti R, Giovannelli A (2018). High-resolution analytical approach to describe the sensitivity of tree environment dependences through stem radial variation. *Forests* 9:134.

Deslauriers A, Rossi S, Anfodillo T (2007). Dendrometer and intra-annual tree growth: what kind of information can be inferred? *Dendrochronologia* 25:113–124.

De Swaef T, De Schepper V, Vandegehuchte MW, Steppe K (2015). Stem diameter variations as a versatile research tool in eco-physiology. *Tree Physiol* 35:1047–1061.

Donnellan Barraclough A, Zweifel R, Cusens J, Leuzinger S (2019). Disentangling the net: concomitant xylem and over-bark size measurements reveal the phloem-generated turgor signal behind daytime stem swelling in the mangrove *Avicennia marina*. *Functional Plant Biology* 46(5) 393-406 <https://doi.org/10.1071/FP18155>.

Drew DM, Downes GM (2009). The use of precision dendrometers in research on daily stem size and wood property variation: a review. *Dendrochronologia* 27:159–172.

Mencuccini M, Salmon Y, Mitchell P, Hölttä T, Choat B, Meir P, O'Grady A, Tissue D, Zweifel R, Sevanto S, Pfautsch S (2017). An empirical method that separates irreversible stem radial growth from bark water content changes in trees: theory and case studies. *Plant, Cell and Environment* (2017) 40, 290–303.

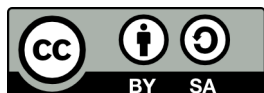
Raffin *et al.* (2021). Améliorer la résistance à la sécheresse du pin maritime. *Forêt-Entreprise*, 256, 14-17.

Sellier D, Segura R (2020). Radial growth anisotropy and temporality in fast-growing temperate conifers. *Annals of Forest Science* 77:85.

Steppe K, Sterck F, Deslauriers A (2015). Diel growth dynamics in tree stems: linking anatomy and ecophysiology. *Trends in Plant Science* June 2015, Vol. 20, No. 6.

Van der Maaten E, Bouriaud O, Van derMaaten-Theunissen M, Mayer H, Spiecker H (2013). Meteorological forcing of day-to-day stem radius variations of beech is highly synchronic on opposing aspects of a valley. *Agric For Meteorol* 181:85–93.

Zweifel R, Haeni M, Buchmann N, Eugster W (2016). Are trees able to grow in periods of stem shrinkage? *New Phytol* 211:839–849.



Cet article est publié sous la licence Creative Commons (CC BY-SA). <https://creativecommons.org/licenses/by-sa/4.0/>.

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « NOV'AE », la date de sa publication et son URL.