

Annex 7. Link van de verschillende databanken en modellen

Eén van de grote uitdagingen van het SimForTree project is de omvang van zowel de input data als van de verwachte simulatiegegevens. De uiteindelijke simulaties werden uitgevoerd voor 4 soorten, 3 klimaatregio's, 3 klimaatscenario's, 2 (populier) tot 20 (beuk) beheeropties per soort en 36 bodemtypes. Daarenboven moest elk van die opties berekend worden voor een bos dat in 2010 elke leeftijdsklasse kan hebben (en dus aangeplant werd in 1860 tot 2010) en voor aanplantingen tussen 2010 en 2150. Dit resulteerde in 625.000 runs van het ANAFORE model. Bij een gemiddelde duur van 5 minuten per run zou dat op één PC ongeveer 5 jaar in beslag nemen

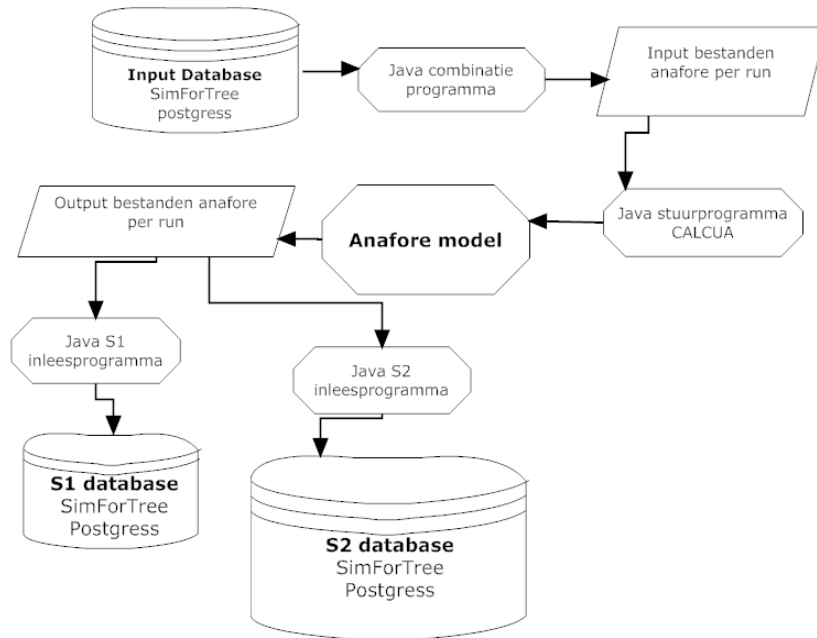
Een aanzienlijke inspanning werd dan ook geleverd om deze simulaties geautomatiseerd te laten verlopen op de supercomputer van de UA (CALCUA). Voor elke run (van de 625.000) zijn invoerbestanden nodig. Het aanleveren van deze invoer diende eveneens te worden geautomatiseerd. Uiteindelijk werd dit in twee stappen bereikt: alle invoer (36 bodems, 50 beheren,...) werd in één invoerdatabase verzameld.

Hiervoor diende enkele belangrijke aanpassingen te gebeuren (cfr. Figuur 1 achteraan):

- De CALCUA is een Linux systeem. Hoewel ANAFORE in Fortran geprogrammeerd is bleek compilering naar Linux toch complicaties met zich mee te brengen, de huidige versie is echter identiek onder Windows en Linux.
- De CALCUA is dermate snel omdat een aantal 'beveiligingen' van Windows achterwege worden gelaten. Daardoor is het systeem echter ook kwetsbaarder. Bv. oneindige lussen lijden tot een crash van het systeem. Het ANAFORE model moest daarom beveiligd worden.
- De snelheid van de CALCUA komt tot stand door het parallel lopen van processen. ANAFORE is een lineair model (het moet van dag to dag doorlopen worden), maar de verschillende runs kunnen wel parallel verlopen. Het verdelen van de runs over de processoren moest echter wel geautomatiseerd worden, waarbij steeds beveiliging werd ingebouwd zodat bij een uitvallen van één processor de juiste runs heropgestart worden. Hiervoor werd een java-stuurprogramma ontwikkeld.
- Een tweede java programma (java-combinatie programma) creëert de juiste combinaties van de invoer files (niet elk beheer moet voor elke soort berekend worden).
- Omwille van de enorme hoeveelheid data, bleek het onmogelijk de volledige versie van ANAFORE te gebruiken. Zo werden de meeste dagwaarden niet behouden, en voor de belangrijkste processen op bestandniveau werden enkel 10-dagelijkse waarden weerhouden. De verkregen data van ANAFORE bestond uit 12 .txt files per run. De CALCUA werd geprogrammeerd om elke run afzonderlijk te bewaren. Een 3^e java-programma (java inleesprogramma) werd ontwikkeld om al deze bestanden in te lezen (met de correcte sleutel om de run te herkennen) in een database (S2 database, zie §4.2 in het SimForTree eindverslag).

- Voor het S1 systeem (zie §4.1 in het SimForTree eindverslag) werd slechts een beperkt deel van de volledige output van ANAFORE gebruikt, deze data werden echter gelijktijdig berekend en afzonderlijk opgeslagen in een aparte database.

ANAFORE - SimForTree: flow of data - frontends



Figuur 1. Schematische voorstelling van ANAFORE flow binnen het SimForTree project.