**Agtergrond vir MasjienLeer gesprek 2018-11-06**

Probe Schedule (PS) maak gebruik van verskeie invoere om grondvogstatus te bepaal en om van daar ‘n schedule te bereken vir die optimale tyd en volume om te besproei.

MasjienLeer kan help om hierdie hele proses deurlopend te ontleed, beherende veranderlikes te bereken en om verfynde vooruitskattings te kan maak.

Historiese metode van schedule berekening:

Eto = Verwysings Evapo-Transpirasie (gemeet as die gebruik van 3” lank geknipte gras).

Eto formule is bereken deur Penman en Monteith jare gelede en bring in berekening weer data  
 naamlik: temperatuur, RH, Wind, son-intensiteit en dampdruk.

Kc = gemete gewasfaktor uitgedruk as faktor van Eto (0.01 – 1.2)

Etc = Gewas se water gebruikstempo bereken vanaf Eto en Kc.   
 Etc = Eto x Kc

Voorbeeld: As Eto=6mm en Kc=0.5, dan is Etc=3mm water gebruik per dag. As die beskikbare water voorraad = 30mm, dan sal die voorraad 10 dae hou mits die weer nie verander nie.  
Die schedule is dus: besproei oor 10 dae env ul aan met 30mm.

PS gaan verder hiermee op 4 maniere:

1. PS berekening hierdie daaglikse model maar bring dan ook gekalibreerde grondvoglesings in vanaf ‘n probe. Die berekende model word aangepas (reggestel) met die werklike lesing om ‘n meer akkurate projeksie te kan maak.
2. PS kyk na die Weather Underground (WU) weervoorspelling vir die volgende 10 dae, trek die weer elemente daaruit en bereken die geprojekteerde Eto met gebruik van die Kimberly-Penman formule.
3. PS se Kc faktore kan met die hand aangepas word om die berekende model te verfyn.
4. Omdat ons die werklike verandering in vogstatus kan meet, kan ons die besproeiings-effektiwiteit ook bepaal en dit aanpas. Bv. Om 33mm in die grond te kry, moet ons 38mm bruto besproei. Die 5mm (38-33) gaan vedamp voor dit in die grond kom.

Die besproeiings effektiwiteit word bereken as:

Eff\_irr = Tot\_irr \* Eff\_factor

Die Eff\_factor hang af van die weer, wind, temperatuur, gront-tipe, en hoe vinning die water die grond penetreer (onder andere). Om die besproeiings-effektiwiteit te bereken is 'n ander veld op sy eie.

Voorbeeld: Eto=6mm maar met ons observasie is Kc=0.45 en met ons probe is beskikbare water dalk 33mm en met ons vooruitskatting van Eto vir die komende 10 dae, is die gebruikstempo vir daardie tyd dalk 3.15mm per dag. Die schedule is nou al baie beter op 10.47 dae (33/3.15) en 38mm.

**MasjienLeer**

Ons het dus hier verskeie veranderlikes wat almal saamwerk om ‘n akkurate schedule te bereken. Masjienleer kan al hierdie veranderlikes volg, ontleed and klassifiseer en terugvoer in ‘n ‘feedback loop’ om Kc en besproeiings-effektitweit te verfyn.

Beste Jac

Myns insiens kan ons die volgende aanpak en ek het hulle gelys in afnemende volgorde van prioriteie, soos wat ek verstaan het:

1. **Voorspelling van Kc** Hierdie is ons **eerste prioriteit**

Identifiseer die periodes waarin **skoon data teenwoordig is**. Maak seker dat Kc bereken work uit data wat nie ge-steur is deur besproeiings, reën of luukse opnames nie.

1. Outomatiese kalibrasie van toerusting Meer ‘n geval van toets dat toerusting reg lees of realistiese waardes gee.
2. Voorspelling van wanneer besproei moet word en hoeveel Hierdie nie nodig nie want as Kc reg is, is dit ‘n maklike som.

'n Paar aspekte wat ek byvoorbeeld sal wil uitklaar is:

* Of daar akkurate historiese Kc waardes beskikbaar is om die algoritme te verfyn? Daar is **nie akkurate Kc beskikbaar nie** – ek maak eintlik die statement dat wat beskikbaar is nie akkuraat is nie en ons wil dit bereken op ‘n nuwe manier.
* Hoe gereeld voorspellings gemaak moet word? Eto en Kc werk op ‘n dagbasis bereken – so **een maal per dag.**
* Hoeveel historiese data in ag geneem kan word in die voorspelling? Ons kan alles gebruik wat daar is wat **in party gevalle 3 jaar is.**
* Wat moet gebeur om die proses as 'n sukses te beskou. Sukses sal wees: Ek wil berekende Kc faktore sien vir uiteenlopende gewasse en weerkondisies wat betroubaar is.

Die faktore wat jy moet in ag neem is die kruks van hierdie effort. Daar is heelwat veranderlikes en **die data kan foutief wees**. Ons moet die veranderlikes deur praat en realistiese norme stel waarbinne jy kan werk. Dan is daar die feit dat een helfte van die data VSA is met duime en die ander metries. Kom ons begin met metries.

Ons stel voor om hierdie te gebruik om die **omskakeling te maak** tussen duim en mm:

* **1" = 25.4mm en 1mm = 0.0393701"**

1. **Eto** wat kom van **weer data** en die **Penman Monteith formule.** Ons moet hierdie as feit aanvaar behalwe as die rou data beskikbaar is en gebrekkig lyk.

Ons sal ook ontbrekende ET0 data moet identifiseer en interpoleer op 'n sinvolle manier.

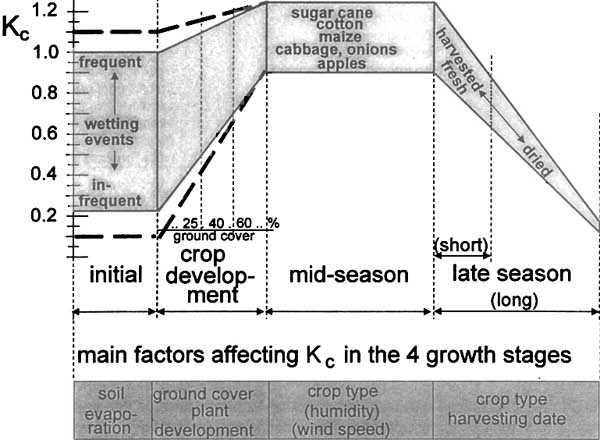
1. **Groeifases** van die gewas. Hierdie het baie te doen met **hitte-eenhede (Growing Degree Days)** <http://msuextension.org/publications/AgandNaturalResources/MT200103AG.pdf>

Kyk ook na

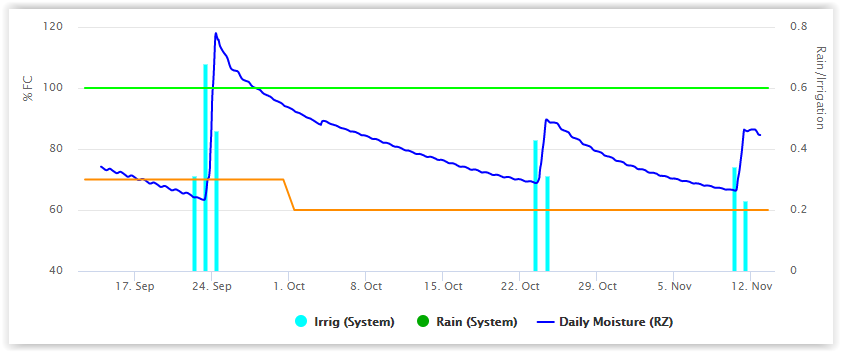
<https://www.researchgate.net/publication/254862120_Impact_of_Crop_Heat_Units_on_Growth_and_Developmental_Physiology_of_Future_Crop_Production_A_Review>

Daar is ook 'n **sterker verband** tussen **GDD en Kc** (as tussen ET0 en Kc).

**Die groeifases van 'n gewas,** volgens die literatuur, is: **(i)** Initial Stage, **(ii)** Developmental Stage, **(iii)** Mid-season, **(iv)** Late season. Sien asseblief die onderstaande grafiek:



Ons sal die **Kc moet verfyn vir al die fases.**



Basiese Reels en Feite

1. Eto, Kc en Etc is **daaglikses syfers**
2. **Eto** is **verwysings-verdamping** gebaseer op gesnyde gras.
3. **Kc** is ‘n faktor van Eto en is ‘n waarde **0.0 <= Kc <= 1.3**
4. **Kc** is verwant aan die **blaarbedekking van ‘n gewas**
5. Die Kc kurwe van ‘n plant is altyd ‘n positiewe klok vorm en maak nooit sporadiese “spikes” nie
6. Eto kom van weerdata en is ‘n waarde **0.0 <= Et0 <= 15mm/dag.** Waardes wat nie binne hierdie interval lê nie moet geflag word as uitskieters en moet nie gebruik word in die berekening van die nuwe Kc nie.

Is al die historiese Weather Underground data gestoor en kan dit gebruik word vir modellering?

Ja, die weerdata word alles gestoor en dit is nie alles WU nie. In **SA** is meeste van **iLeaf** maar ons kan die bron sien.

1. As die Eto te hoog raak vir die plant om te transpireer maak die huidmondjies toe om die plant te beskerm en geld die Etc reel nie meer direk nie. Eto perke hang af van gewas maar ‘n algemene reel is dat 12<-->15mm die maksimum is. Na daardie punt val die Etc nie na 0 nie *maar is laer* as wat die verband Etc=Eto\*Kc voorspel. Alle waardes van ET0 wat buite die interval [0, 15 mm/dag] lê is nie geldige data nie. Die rede hoekom ET0 > 15mm/dag kan wees is as gevolg van **foutiewe weerstasie data**: ons sal moet uitkyk vir sulke foutiewe data in ons datastelle.
2. Etc = Eto\*Kc

Hierbo staan Etc = Eto x Kc, dus Kc=Etc/Eto? ~~Of geld Etc = Eto/Kc as die Eto te hoog raak?~~

Jou eerste stelling is reg. Kc = Etc/Eto met die perke soos bespreek (met ander woorde wanneer 0mm/dag <= ET0 <= 15mm/dag).

1. ‘n plant se groeifases hou direk verband met Hitte-eenhede of GGD (Growing Degree Days)
2. Daarom is Kc nie direk gekoppel aan datums nie maar wel aan GGD (wat min of meer met datums ooreenstem. Daar is 'n **sterker** verband tussen **GDD en Kc** (as tussen, ET0 en Kc).
3. As die grondprofiel oorvol is, is daar **dreinering** + Etc (crop water use)
4. Dreinering is ‘n eienskap van grond tekstuur, kompaksie en beperkende lae onder die grond.

Masjienleer moet dalk so werk:

1. Vir elke gewas-blok warmee jy gaan werk – kyk of daar weerdata is waarmee GDD bereken kan word.

Vir 'n beginpunt gaan ons werk met Golden Delicious Appel Data. Die ge-assosieerde Probe-Ids is as volg: {370, 371, 372, 384, 391, 392, 891, 1426}. Probes 370–891 is by Kromfontein-Ouplaas; Probe 1426 is by Beau-Haru.

1. Kyk of die weerdata volledig is. Indien sporadiese happe voorkom, interpoleer data sodat GDD berkenen kan word. Et0 moet ook beskikbaar wees. Interpoleer ontbrekende ET0 waardes op 'n logiese en sinvolle manier.
2. Kyk of probe data volledig is en vul gapings in as dit klein is en die interpretasie logies is. Die data wat hierna verwys word is die grondvogstatus soos deur die probes gemeet.
3. (Jy sal moet aanvaar dat die kalibrasie reg is)
4. In PS het ons ‘n bestuurs-wortelsone en die **totale wortelsone**. Vir hierdie doel stel ons belang in die **totale wortelsone** want ons wil die ***totale wortelaktiwiteit* in ag neem.**

Metodiek:

1. Werk met blokke op ‘n gegewe plaas vir die begin.
2. Kry weerstasie naam gekoppel aan plaas.
3. Kry weerdata vir die plaas.
4. Evalueer weerdata en bereken GDD en **maak beide GDD en Eto skoon.**
5. Kry blok naam en gewastiepe.
6. Trek Etcp, reen en besproeiing uit WaterBalans tabel
7. Evalueer Etcp en maak skoon om **Reen & besproeiing uit te skakel**

Berekenings:

1. Vir **SA** begin GDD berekening by **1 Augustus** en bereken

**GDD = (Tmin+Tmax)/2 – 10 C**

In die bostaande vergelyking is die **'Temperature Base' 10C** gemaak. Volgens internet bron: *"****Temperature Base*** *is the temperature below which plant development stops."* Vir appel gewasse is 'n Temperature Base van 10C voldoende.

1. Gaan haal Etcp uit waterbalans tabel

Verwys Etcp na die Etc wat gemeet ~~of beraam~~ is? Etcp is my eie begrip maar beteken die Etc **soos gemeet met die gekalibreerde probe**. Jou masjienleer is juis om te toets vir **goeie Etcp** and dan die Kc darmee te verfyn sodat ons volgende beraming van Etc = Eto\*Kc beter is.

ET\_cp is die **verskil tussen opeenvolgende "Profile" lesings** (Die "profile" lesings word gevind in die waterbalans tabel).

1. Skakel dae uit waar daar **bevestigde reen of besproeiing is** sodat ons net **suiwer Etc** sien.
2. **Interpoleer happe** in Etc om gladde kurwe te kry
3. Ondersoek verband tussen **Eto, GDD en Etcp**
4. As die verband ‘n patroon vorm, bereken **Kcp = Etcp/Eto.**
5. Kcp moet ook ‘n **gladde kurwe wees** met waardes in die **0.0 – 1.3 range.**
6. **Stoor die daaglikse Kcp** in ‘n Kcp table saam met blok naam, Crop Type, Eto, Etcp en GDD
7. Sodra daar genoeg van elke Crop type is kan ons die syfers begin vergelyk en soek vir die rede vir afwykings. Dit kan wees foutiewe data, foutiewe kalibrasie, geografiese area, foutiewe gewastiepe ens.

Finale produk:

In hierdie hele proses moet masjienleer die beginsels gevoer word en norme gedefinieer word sodat ons eindig met ‘n slim metode om Kcp te bereken gebaseer op bestaande data en om Kcp en Etc te voorspel gebaseer op al die versamelde data.

Doen herberekenings om die logika te laat werk in VSA se duime.

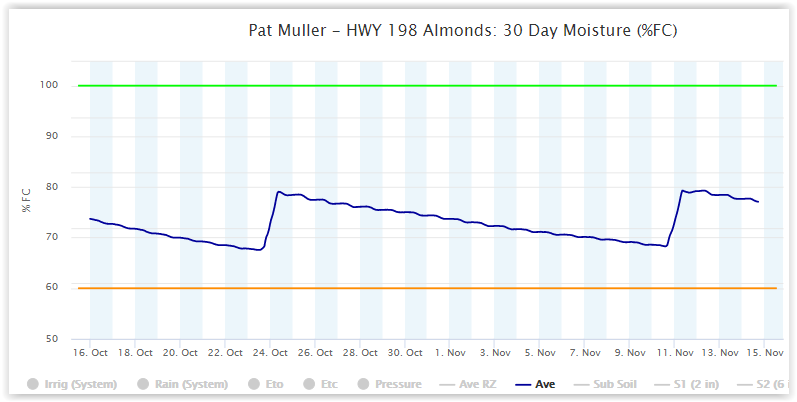
Beskikbare data:

Die grafiek onder is seker die skoonste data wat mens sal sien.

Ons is besig om die algoritme in plek te sit om Etcp op bestaande data te bereken. Andre sal dan die proses afskop om all blokke te bereken vanaf 1 Augustus. Vir **VSA** sal dit wees vanaf **1 April.**

Ons sal dus hierdie Etcp kan gebruik om die Kc te help beraam? Korrek, maar jy moet **eers Etcp goed ondersoek**. Met ander woorde, maak seker jy gebruik ETcp waardes wat nie ge-affekteer is deur reën, besproeiings en luukse opnames nie.

Kan ons dit reeds handmatig doen as dit ten tyde van die modellering nog nie gedoen is nie?  
Ja, ons kan teruggaan en enige met hand bereken.



Kommentaar:

Met masjienleer leer ons gewoonlik ons modelle om 'n seker waarde, wat ons weet korrek is, te voorspel.

In hierdie geval is dit die Kc, maar die probleem is dat die Kc nie noodwendig akkuraat was nie.

Die model gaan dus nie leer van die Kc wat korrek is nie, maar gaan die **Kc bereken d.m.v. ‘n formule.**

Hierdie formule sal dan gereeld aangepas word, gebaseer op die beskikbare data.

Is dit reg so?

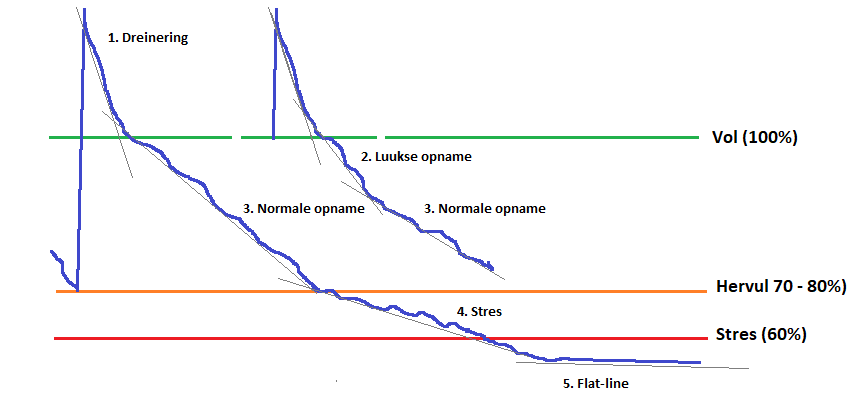
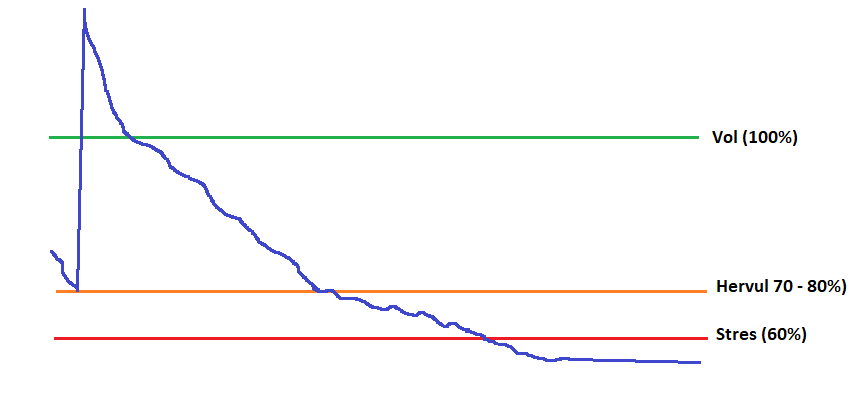
Ja, jy is min of meer reg. In hierdie geval moet ons **aanvaar die Etcp is korrek**. Maar ek dink ons moet eers net met Etcp werk en seker maak ons verstaan dit en kan dit korrek voorspel onder hierdie omstandighede: (hier is waarskynlik duplikate van wat bo gelys is maar miskien iets nuuts ook)

* Etcp is op sy beste as die boom **gesond is** (95% van die tyd), en die watervlak is tussen 80 – 100%
* As die boom stres van te min water (droogte-stres, onder 70%) dan is die opname minder. Dit is waar vir die situasie maar ons moet nie die Kc hierop baseer nie want as dieselfde boom nie gestres is nie sal dit meer gebruik onder die identiese omstandighede.
* As die profiel oorvol is (105%+) dan is daar min suurstof in die grond en is die plant in ‘n nat-stres sistuasie. Ons moet hierdie ook optel en flag.

So ons moet afwykings later probeer identifiseer en probeer bepaal waneer dit as gevolg van een van die tipes stres is? Jy moet vir hierdie oefening eers die **normale gebruik identifiseer en op dit werk.** Dit behoort te wees in die **100 – 70% sone of 90 – 70% sone.**

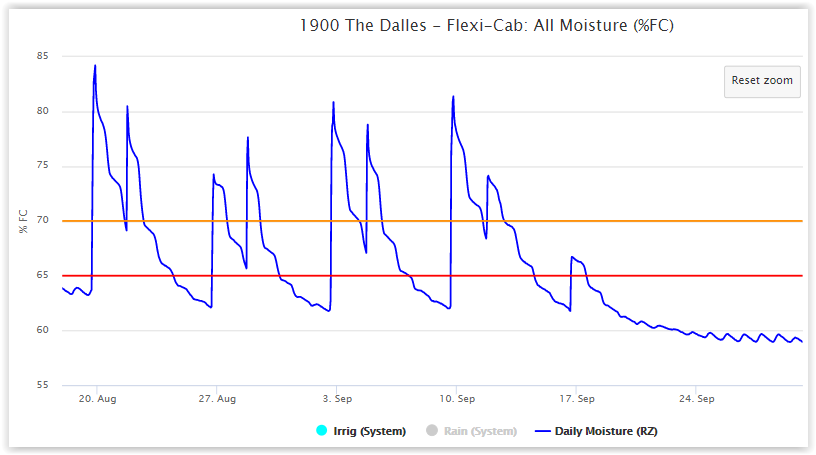
* As die profiel oorvol is en die grond kan goed dreineer (veral sand) dan val die waterinhoud vinnig tot by veldkapasiteit (VK) maar dit is dan **Etc + Dreinering**. Ons moduleer dreinering gebaseer op grondtiepe. Vir hierdie doel moet jy kyk na Vogstatus (105%+), grondtiepe en ook Etc vergelyk met aanvaarde norme. ‘**n Boom gaan nie oor Kc = 1.1 nie.**
* Ek kom telkens terug na **Kc = Etc/Eto** wat die bestaande geloof is maar as jou werk dit verkeerd bewys – kom ons uit met iets revolusiner !! Ek het reeds sekere goed verkeerd bewys.
* As die boom onder stres was en nou water kry, is die opname vir die eerste 2 dae meer as normaal. (Dink aan ‘n mens wat gedehereer is na heel dag werk in die son – jy drink meer om in te haal en gaan dan weer normaal aan.
* Ek dink daar is 5 fases wat ons kan onderskei in die water gebruik patroon. Kyk na die 2 figure onder. Op die 2e het ek die fases ingetrek en genommer. **On wil fase 3 (normale Etcp) probeer definieer in die data.** Onder is die fases weeer beskryf vollediheidshalwe:

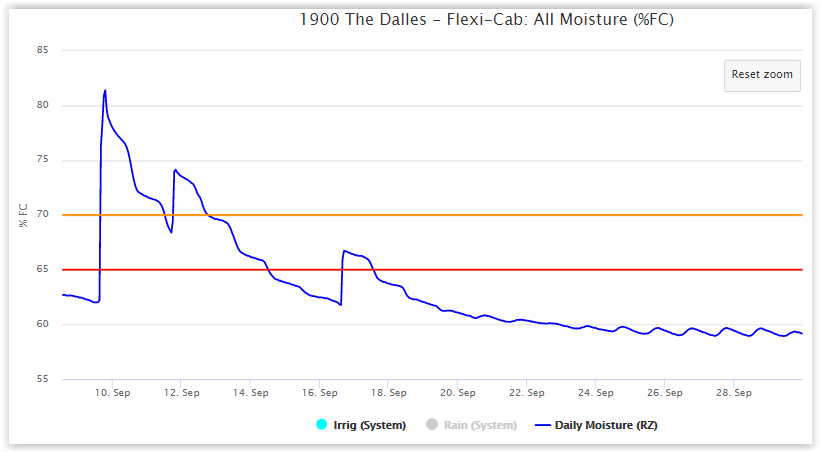
1. **Dreining + Etc**. Dreinering hang nie net af van die grond nie maar ook van hoe vol die ondergrond is. As die profiel bestaan uit sand op ‘n klip laag gaan die baie goed dreineer totdat die water begin opdam op die klip dan neem dreinering af tot omtrent 0 mm/dag. Hierdie fase is ‘n “wildcard” en ons wil sekere wees ons neem dit nie in ag nie.
2. **Luukse opname**. Hier sien mens as die gewas droog was en nou water kry. Dit hou gewoontlik aan vir so 2 dae dan gaan alles na normaal.
3. **Normale opname.** Hierdie is die Etcp wat vir ons die standaard Kc moet gee.
4. **Stres**. My skets wys dit nie so nie maar hierdie is eintlik ‘n lyn wat asemtoties na 0 Etc gaan.
5. **Flat-line**. Daar is nie meer merkbare opname in die profile nie en mens wonder hoe die plant aan die lewe bly.



Kyk na die **2 grafieke hier onder**. Dit is my eie Cabernet wingerd – en ja die wyn is lekker!

Wyndruiwe word onder beheerde stres gehou en ek maak nooit vol nie. Ek gee 2 kort besproeiings per week (2 x 4uur want my timer kan nie meer as 4 uur gee nie – anders sou ek 1 x 6 uur gedoen het). Dinge gebeur vinnig maar jy kan die fases sien: Lukkse gebruik vir 1 dag, Normaal vir 2 dae, Stes vir 2 dae dan gee ek weer water. Die Flat-line het aangehou tot 30 September, to is die Brix by 26 een toe oes ek. Terloops, wat nie wys op hierdie grafiek nie, die wingerd was vir 2 maande in Flat-line toestand en ek het dit geen water gegee nie tot die groei gestop het.





Aangeheg is ‘n spread sheet met lyste van blokke met skoon data waarmee jy kan werk. Ek sal hierdie lys aanhou uitbou en nou moet ek sommer begin onderskei tussen gewasse wat jy dan apart moet hanteer.

Baie dankie, sal jy asseblief bevestig wat nog bygevoeg gaan word vir fase 1?

Ek sal vir jou nog skoon data soek om die algoritme mee te bou – ek neem aan meer is beter. Maar ek verwag nie dat ons al die gewasse moet deurwerk nie. Ek wil vir jou **genoeg gee om die beginsels mee te toets.** Dan as ons weet die werk, kan ons uitbrei na ander kultivars toe.

Ek stel voor jy begin met appels en los die pere en kersies vir later en vir kruis-toets.

Hieronder is van die veranderlikes wat ek geïdentifiseer het wat ons kan gebruik.

Dit sal help om daaglikse lesings vir **elke dag te kan kry, vir die laaste drie jaar**, indien moontlik, of hou wou jy dit doen?

Ek weet nie of ek goeie data het wat so oud is nie. Ons sisteem is nie vriendelik om so ver terug te gaanie. Ek sal dit met Andre ook opneem.

* Eto: Vir blokke/plase wat data trek van dieselfde weerstasie verwag ons dat die Et0 lesings **identies moet wees** vir die geassosieerde probes.
* Etc: Hierdie lesings **kan identies wees** vir verskillende probes. Dit is 'n teoretiese "crop water usage" gebaseer op Kc wat deur Jac verskaf word.
* Etcp: Vir verskillende probes kan die lesings van Et\_cp **NIE identies wees nie**, maar hulle kan wel naby aan mekaar wees. As dit wel identies is, moet ons met André praat want hy het heel waarskynlik een probe se data herhaal.
* Penman Monteith formule se waarde die waarde is **Eto**
* Kc – oorspronklik
* Kc – aangepas
* Reën ons wil reen eers miskyk en probeer data identifiseer met geen reen of besproeiing nie
* Besproeiing – soos vir reen want hierdie is nog ‘n veranderlike
* Grondvoglesing vanaf probe – jy behoort die nie nodig te kry nie want ons het reeds die probe lesing gekalibreer na %veldkapasiteit en die **mm lesing van wortelsone en profile in die WB tabel gesit per dag.** Die probe data is rou lesings en beteken min op sigself.
* Weather Underground (WU) weervoorspelling of ander weerdata – ons gaan nie met voorspellings nou werk nie maar kyk na **historiese data**
* Weerstasie naam gekoppel aan plaas
* Besproeiingseffektiwiteit – sal later in die modulering ingebring word
* Gewas
* Kultivar
* Plaas
* Blok
* Land – blok en land is maar dieselfde ding
* Indikasie of lesings metriek is of ‘imperialisties’ – baie belangrik ! ek gee eers vir jou net metries om mee te werk
* Growing Degree Days (GDD)
* Blaarbedekking – ek het nie hierdie parameter nie maar is **gekoppel aan GGD en Kc**
* Grondprofiel
* Is die grondprofiel oorvol of nie
* Dreinering
* Grondtekstuur
* Kompaksie
* Beperkende lae
* Ander probe data – ander probe data wat ek baie van het is **neutron data** wat net **een maal per week** gelees word.
* **Is die probe gekalibreer of nie? ons kan net met gekalibreerde data werk.**
* Bestuurs-wortelsone
* **Totale wortelsone**
* Totale wortelaktiwiteit
* Waterverbruikfase
* Algoritme om Etcp op bestaande data te bereken, se berekende waardes of die logika daarvan- ons doen hierdie reeds – jy kan die logika met Andre bespreek maar jy hoef dit nie te bereken nie.
* Die Average Root Zone Moisture (azm) is veronderstel om net die wortelsone te verteenwoordig.
* Dit is altyd geldig dat die lesings van **Profile >= azm** (Average root zone).

Voorstel kommentaar:

* ek het **appels gekies om mee te begin** net omdat ons verwag om mooi skoon data te kry wat naby mekaar val maar die logika moet kan werk met enige gewas.

Reg so

* ek wil he daar moet 'n veld geskep word in die Kc databasis waar die ge-interpoleerde Kc gepak word. Vandaar wil ek die grafiek evalueer teen die Eto en GGD en as dit goed lyk, moet dit die permanente rekord vervang. Ons sal vir Andre moet intrek om die velde te skep.

Is dit ‘n grafiek wat tans bestaan?

* verder wil ek weet waar die finale produk sit, hoe ek dit aktiveer en hoe die kruiskontrole lyk

Voorstel? Dalk ‘n ekstra veld?

* dit moet kan werk op US data ook. Die formule bly Etc = Eto \* Kc. Dit is net mm wat duime word.

Reg so

* gaan die produk eindig as 'n model met veranderlikes wat ek kan verstel om goed te laat pas?

Ja

