AMICAF/MOSAICC seminar
ITCAF, Manila, 13-14 Sept. 2012

A general introduction to crop models

René Gommes
rene.gommes@agro-impact.com

Overview

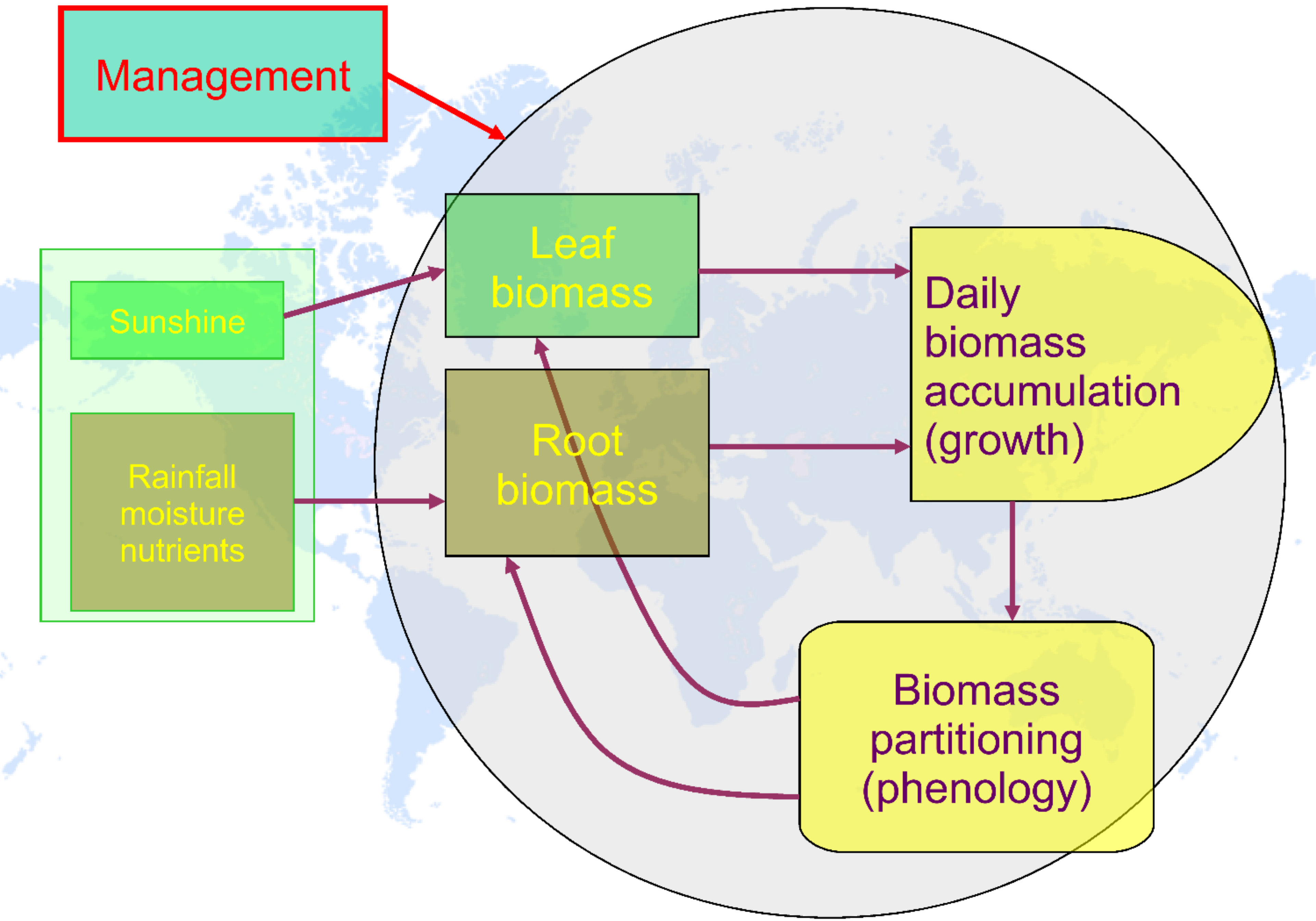
- There are models, models, models, models and models
- Turning a bit philosophical: some issues, considerations & conclusions

Models, models models, models and models

- Process-oriented simulation models
- Statistical
- Non-parametric
- FAO AgroMetShell, the “ancestor” of wabal
- “Other”

Simulation models: the scope

- Realistically & “scientifically” mimick actual physiological mechanisms and interactions of plant & environment, incl. management
- Accurate and versatile
- calibration switches tweak model into good **qualitative** fit to reality
- Models come in “schools” or “families”: EPIC, CERES, WOFOST (SUCROS, ARID, MACROS, ORYZA1...)


Management
Sunshine
Rainfall moisture nutrients
Leaf biomass
Root biomass
Daily biomass accumulation (growth)
Biomass partitioning (phenology)

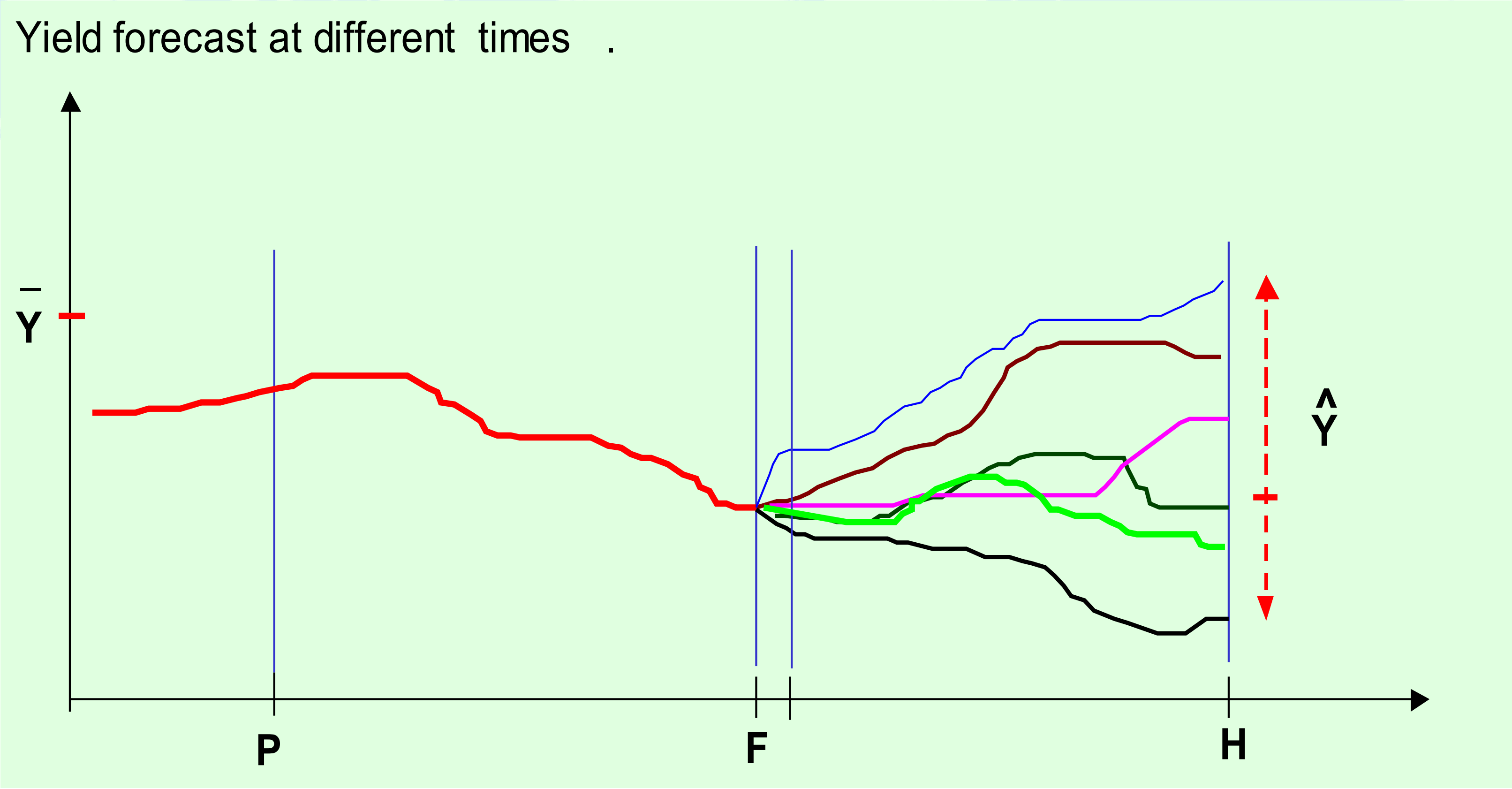
A word about the "inner workings" of models: variables, parameters, inputs...

- State variables (global): completely describe the state of the "system", e.g. biomass on day d in g m^{-2})
- Rates: the speed at which the state variables change (e.g. rate of change of biomass in g $m^{-2} d^{-1}$
- Derived variables: computed from state variables (less fundamental nature, e.g. Leaf Area Index, in m^2 of leaves per m^2 land area, computed from biomass)
- Parameters or "switches": constants that describe links between variables and rates
- Input variables: measure external action on the system (e.g weather & management)

More the "inner workings"

- Models contain many ad hoc functions; they are often less "scientific" than assumed
- In practice, models are computer programmes
- Time step is usually daily
- Many models compute various "biomasses": water limited (B_W), energy limited (B_E) and nutrient limited (B_N) and adopt $\min(B_W, B_E, B_N)$ as final yield proxy
- The "budget nightmare": energy-water, nutrients, biomass

Using a model for forecasting: past and future weather



Simulation models: typical components

- Biomass accumulation (assimilation)
- Phenology (or development) and biomass partitioning (incl. Respiration and root development)
- Nutrient budget
- Soil & plant water budget

Photosynthesis: orders of magnitude

- $6CO_2 + 12H_2O \rightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O$ – 2880 KJ mol^{-1} sugar (1 mol = 180 g)
- 22g CO_2 + 18 g H_2O →

  15 g $C_6H_{12}O_6$ + 16g O_2 + 9 H_2O
- [CO_2] ~400 ppmv ~775 mg/m^3 ~0.75 g/m^3
- 1 m^2 of leaf can produce 1 g of sugar in 1 hour, req. 22/15 ~ 1.5 g CO_2
- 1.5 g CO_2 is the amount contained in 2 m^3 of air

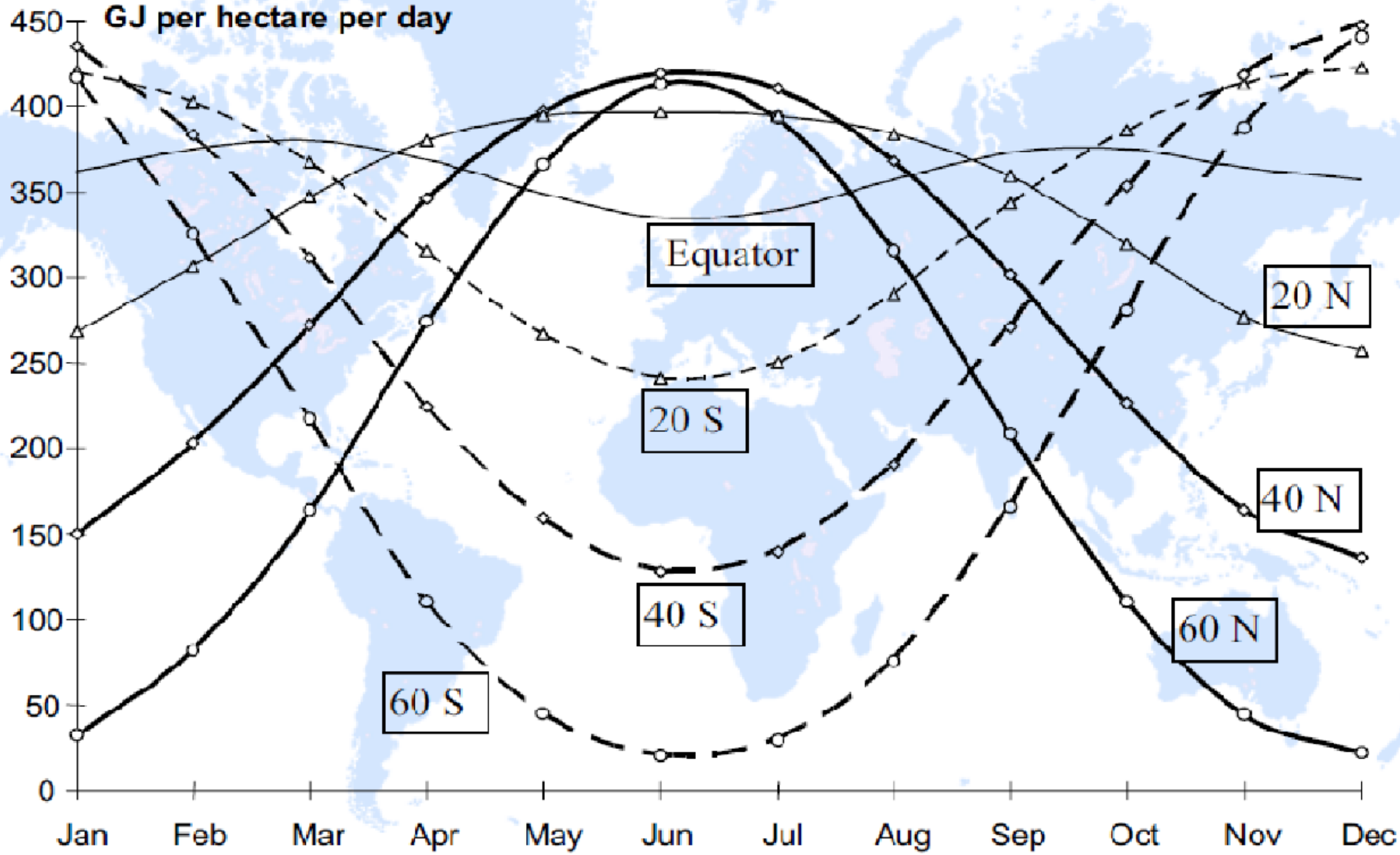
Photosynthesis: potential biomass (Monteith)

- $DM = H \cdot Eff_H \cdot Eff_a \cdot Eff_c$
  - DM = dry matter g ha^{-1} day^{-1}
  - H global net radiation J ha^{-1} day^{-1} ~ 50% of extraterrestrial radiation
  - Eff_H, fraction of H which is PAR (~ 0.33-0.50)
  - Eff_a, interception efficiency, f(LAI, geometry... ~ 0.33 to 0.50)
  - Eff_c, conversion efficiency, 2 g DM/MJ for C3 plants (3 for C4)

Photosynthesis: overall efficiency

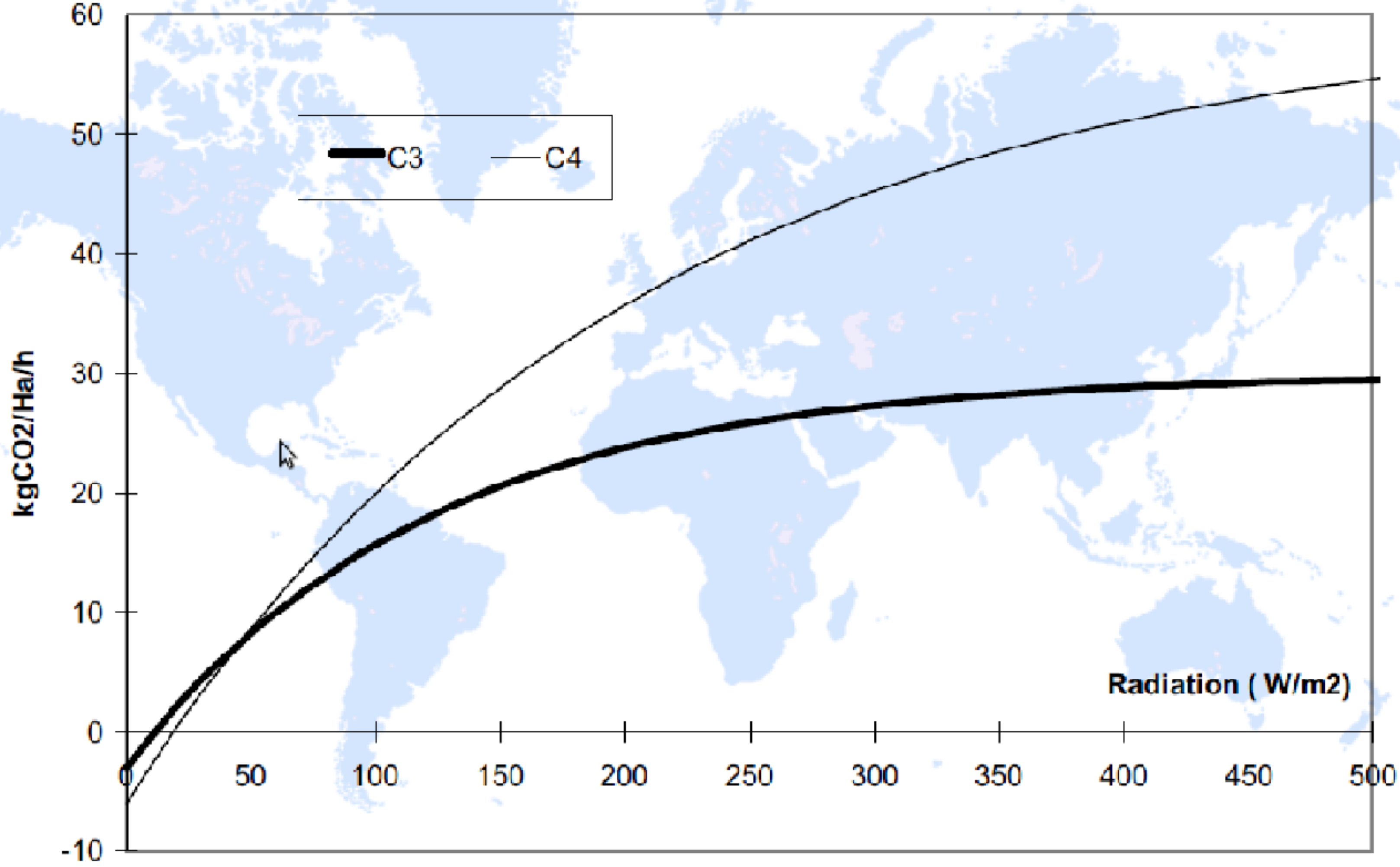
- Compare
  - Chemistry: 1 MJ yields 60 g sugar (2880 kJ ~ 3 MJ / 180 g)
  - Biology: 1 MJ yields 2 g DM
- This is because (and/or) biomass is not just "accumulated"
  - About 50% of sugar is consumed in dark and photorespiration: maintain the plant structure
  - DM is not only sugar, but also "value added" fats, starch, cellulose, proteins...
  - Maintenance respiration (maintain the "structure")
  - Develop & live (grow roots, flowers, attract insects, repair damage...)
- Net efficiency of a leaf is about 5% of incident radiation.
- Canopy Incident sunlight to biomass efficiency: from 1 % (typical crop) to 8% (sugarcane). Most plants store 0.25-0.50% in the product (grain...)

Exercise...



GJ per hectare per day
450
400
350
300
250
200
150
100
50
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Equator
20 N
20 S
40 N
40 S
60 N
60 S

Photosynthesis: function of absorbed PAR (Ha=Ha of leaf)



Photosynthesis: the equations

$$F_n = F_d + (F_m - F_d)\left(1 - \exp\left(-\frac{E_{lc} . R_{HC}}{F_m}\right)\right)$$

			C3-plants	C4-plants
Net assimilation	Kg CO_2 / Ha leaf / hour	F_n		
Maximum rate of net assimilation	Kg CO_2 / Ha leaf / hour	F_m	30 (15 to 50)	60 (30 to 90)
Net assimilation in the dark	Kg CO_2 / Ha leaf / hour	F_d	-3	-6
Absorbed radiant flux in the 400-700 nm range	joule / m^2 / s	R_{HC}		
Efficiency at light comp. point	Kg $C0_2$ / Joule	E_{lc}	0.25	0.30
Temperature-dependent F_m ?			No	Yes

Simulation models: typical components

- Biomass accumulation (assimilation)
- Phenology (or development) and biomass partitioning (incl. Respiration and root development)
- Nutrient budget
- Soil & plant water budget

Phenology (1/2)

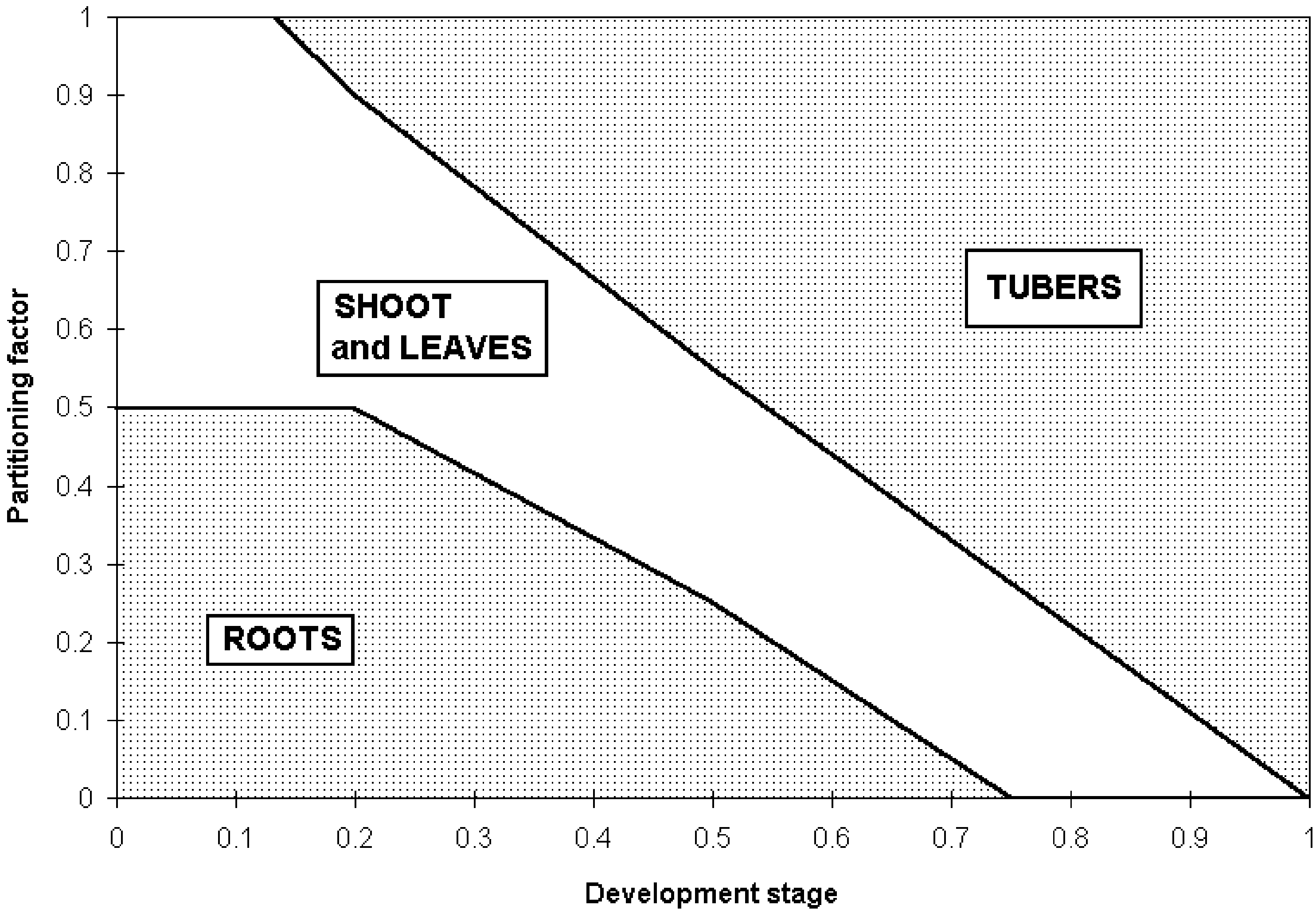
- Typically one of the most empirical components of models
- Growing Degree-Days (GDD) or Sums of Degree-Days "cannot go wrong"
- GDD miss all qualitative effects (more suitable for climate where heat is limiting)

$$GDD_S = \sum_{\text{Planting day}}^{\text{Day on which stage } S \text{ is reached}} (T - T_b) \quad where \quad \begin{array}{c} T - T_b \text{ is taken as 0 when } T < T_b \\ T \text{ is taken as } T_u \text{ when } T > T_u \end{array}$$

Sample GDD from *Wikipedia*

Common name	Latin name	Number of growing degree days baseline 10 °C
Dry beans	Phaseolus vulgaris	1100-1300 GDD to maturity depending on cultivar and soil conditions
Barley	Hordeum vulgare	125-162 GDD to emergence and 1290-1540 GDD to maturity
Sugar Beet	Beta vulgaris	130 GDD to emergence and 1400-1500 GDD to maturity
Wheat (Hard Red)	Triticum aestivum	143-178 GDD to emergence and 1550-1680 GDD to maturity
Oats	Avena sativa	1500-1750 GDD to maturity
European Corn Borer		207 - Emergence of first spring moths
Corn (maize)	Zea mays	2700 GDD to crop maturity
Forsythia	Forsythia spp.	begin flowering at 1-27 GDD
Common lilac	Syringa vulgaris	begin flowering at 80-110 GDD
Red maple	Acer rubrum	begins flowering at 1-27 GDD
Black locust	Robinia pseudoacacia	begins flowering at 140-160 GDD
Purple loosestrife	Lythrum salicaria	begins flowering at 400-450 GDD
Sumac	Rhus typhina	begins flowering at 450-500 GDD
Butterfly bush	Buddleia davidii	begins flowering at 550-650 GDD

ΔBiomass partitioning



Simulation models: typical components

- Biomass accumulation (assimilation)
- Phenology (or development) and biomass partitioning (incl. Respiration and root development)
- Nutrient budget
- Soil & plant water budget

Nitrogen (as dealt with in CropSyst)

- Proteins contain about 20% of N, i.e. 0.5 and 1.5% of the dry matter. N_{max} is the maximum (reference) crop nitrogen concentration and N_W the actual concentration (both Kg N (Kg DM)$^{-1}$
- "Growth" N demand on day J (in Kg N Ha^{-1})

$$GD = N_{\max} \times \Delta W$$

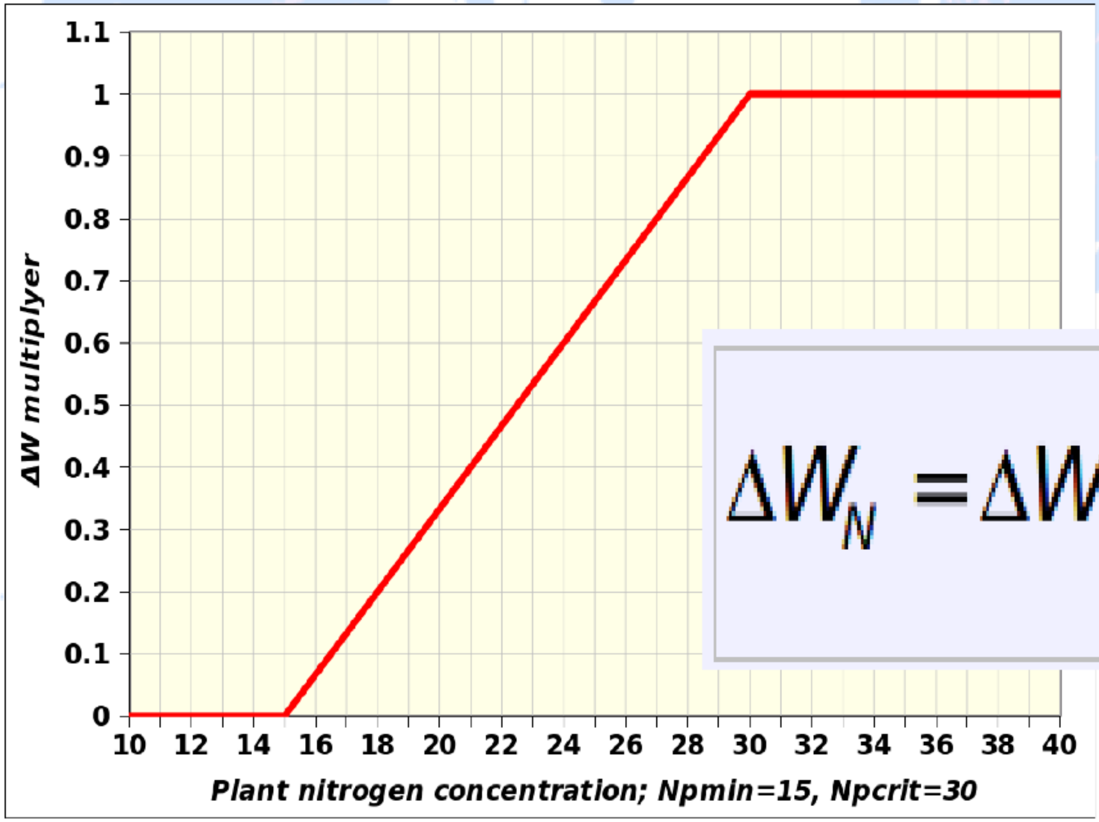
- Demand deriving from the current deficit in the plant at the beginning of J when the biomass is W (Kg DM Ha^{-1})

$$DD = W(N_{\max} - N_W)$$

- Total demand

$$TD = DD + GD$$

Nitrogen-limited daily biomass accumulation (CropSyst)



ΔW is the water and radiation limited growth

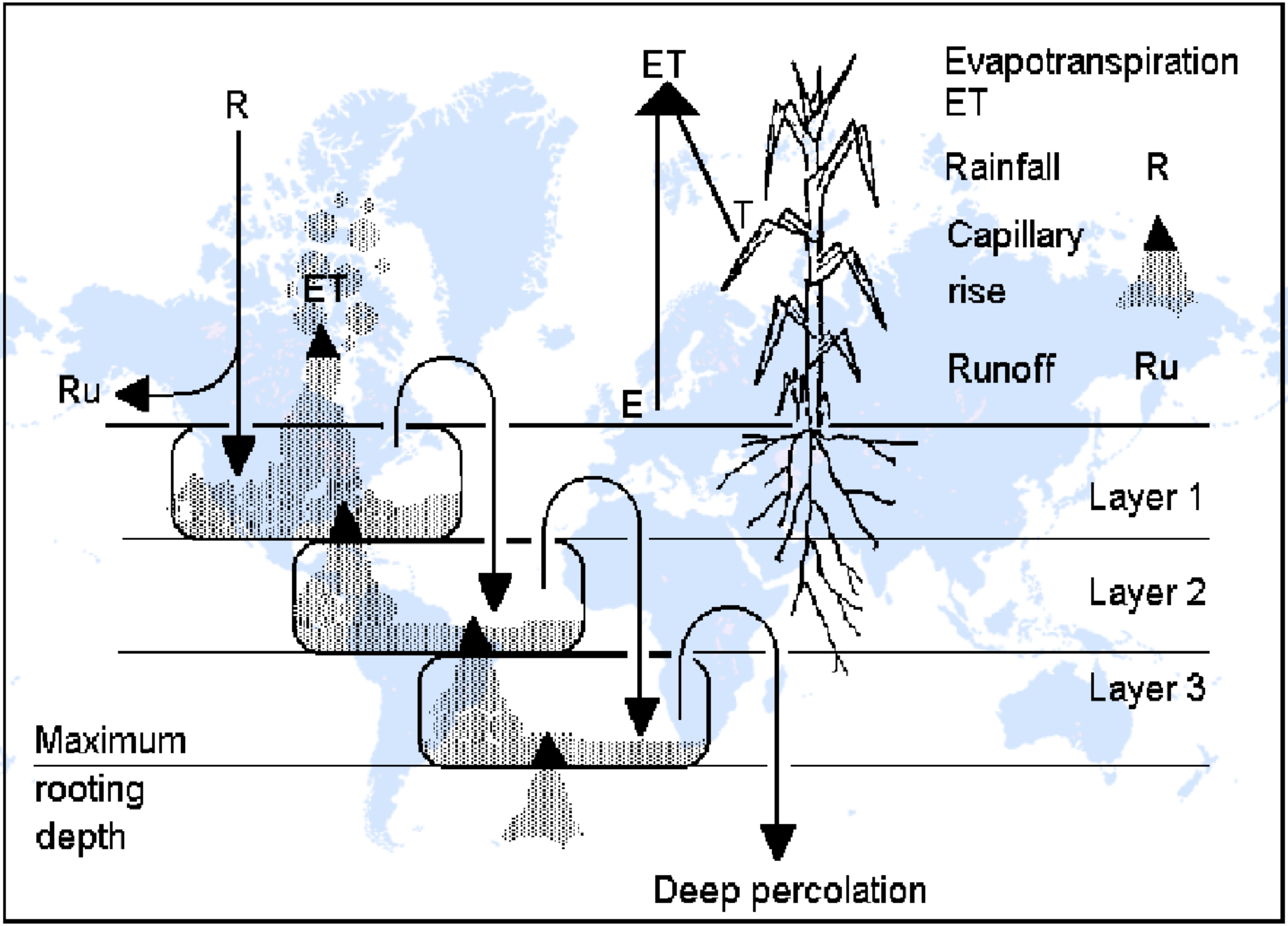
$$\Delta W_N = \Delta W \left(1 - \frac{N_{pcrit} - N_p}{N_{pcrit} - N_{pmin}}\right)$$

Simulation models: typical components

- Biomass accumulation (assimilation)
- Phenology (or development) and biomass partitioning (incl. Respiration and root development)
- Nutrient budget
- Soil & plant water budget

Water budget

- Water supply depends on the amount of water (rainfall, irrigation) that enters the soil
- Soil water availability depends on the balance between the strength with which water is held in the soil and the strength of the demand exerted by the plant
- Many different levels of complexity of soil water budget



ET
R
T
ET
Ru
E
Evapotranspiration ET
Rainfall R
Capillary rise
Runoff Ru
Layer 1
Layer 2
Layer 3
Maximum rooting depth
Deep percolation

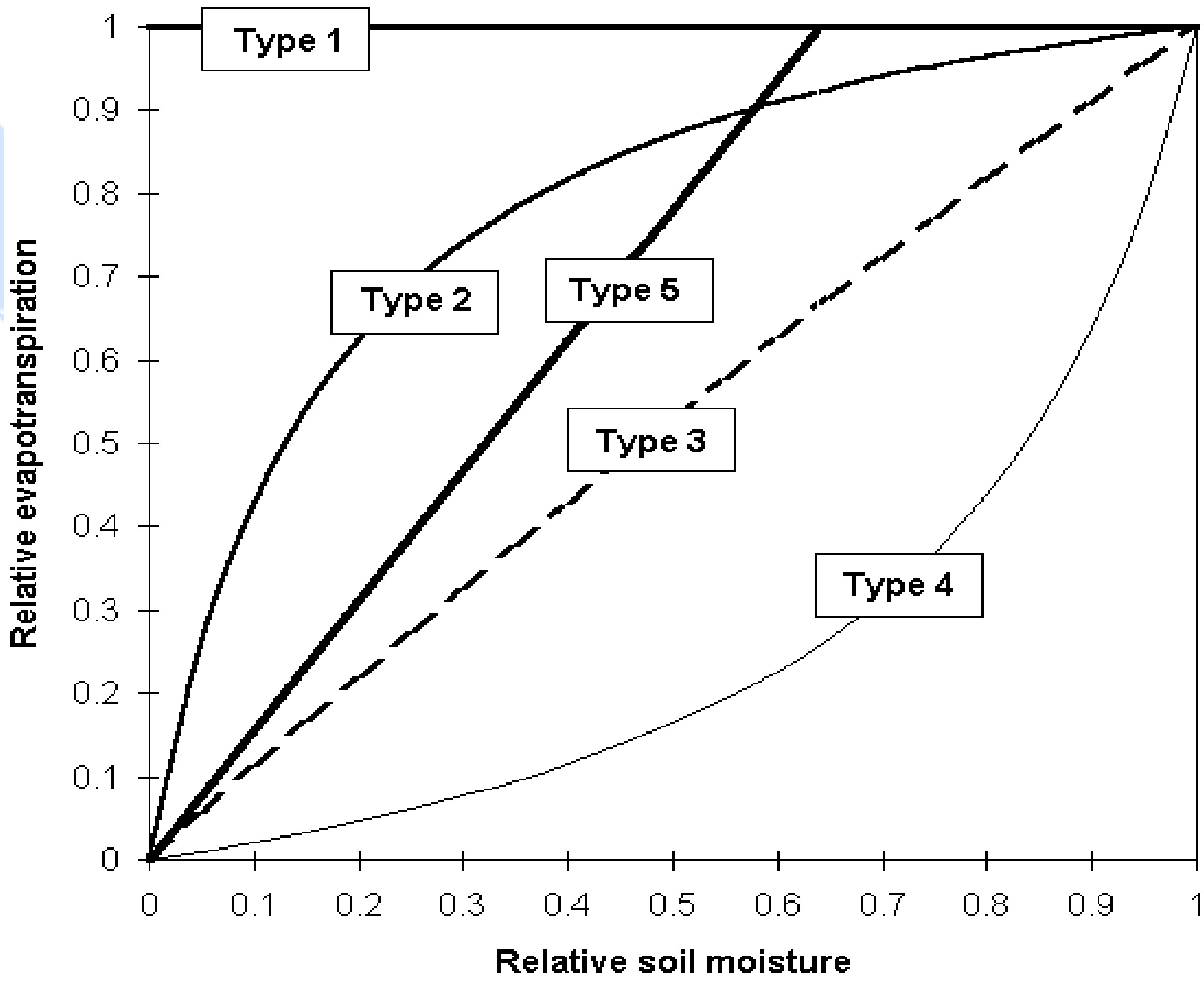
- Main driving force for water demand is the evaporative demand of the atmosphere, measured by the evapotranspiration potential (ETP)
- ET is often partitioned between E and T (for instance):

E = ET – T

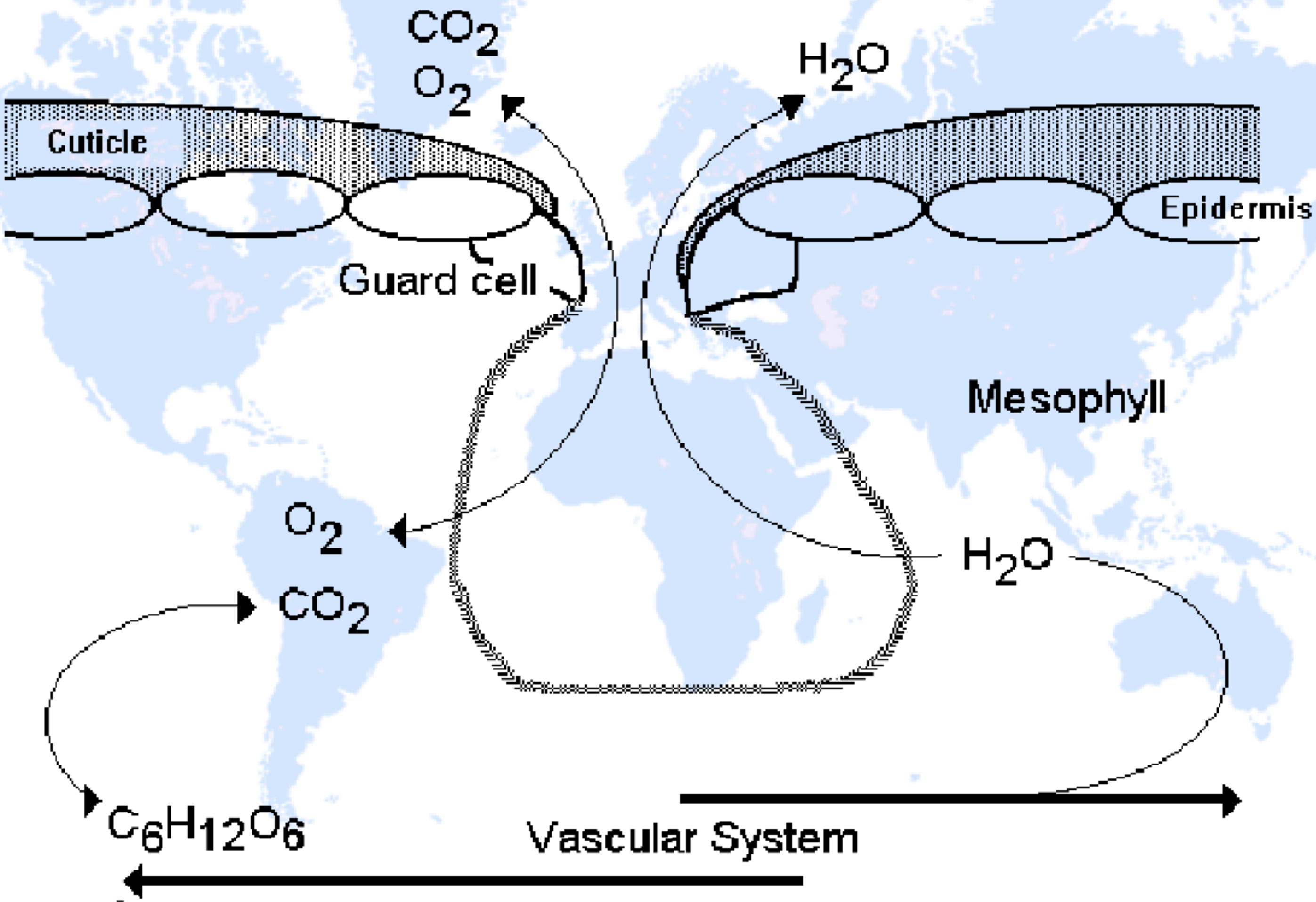
T = LAI * ET if LAI<1

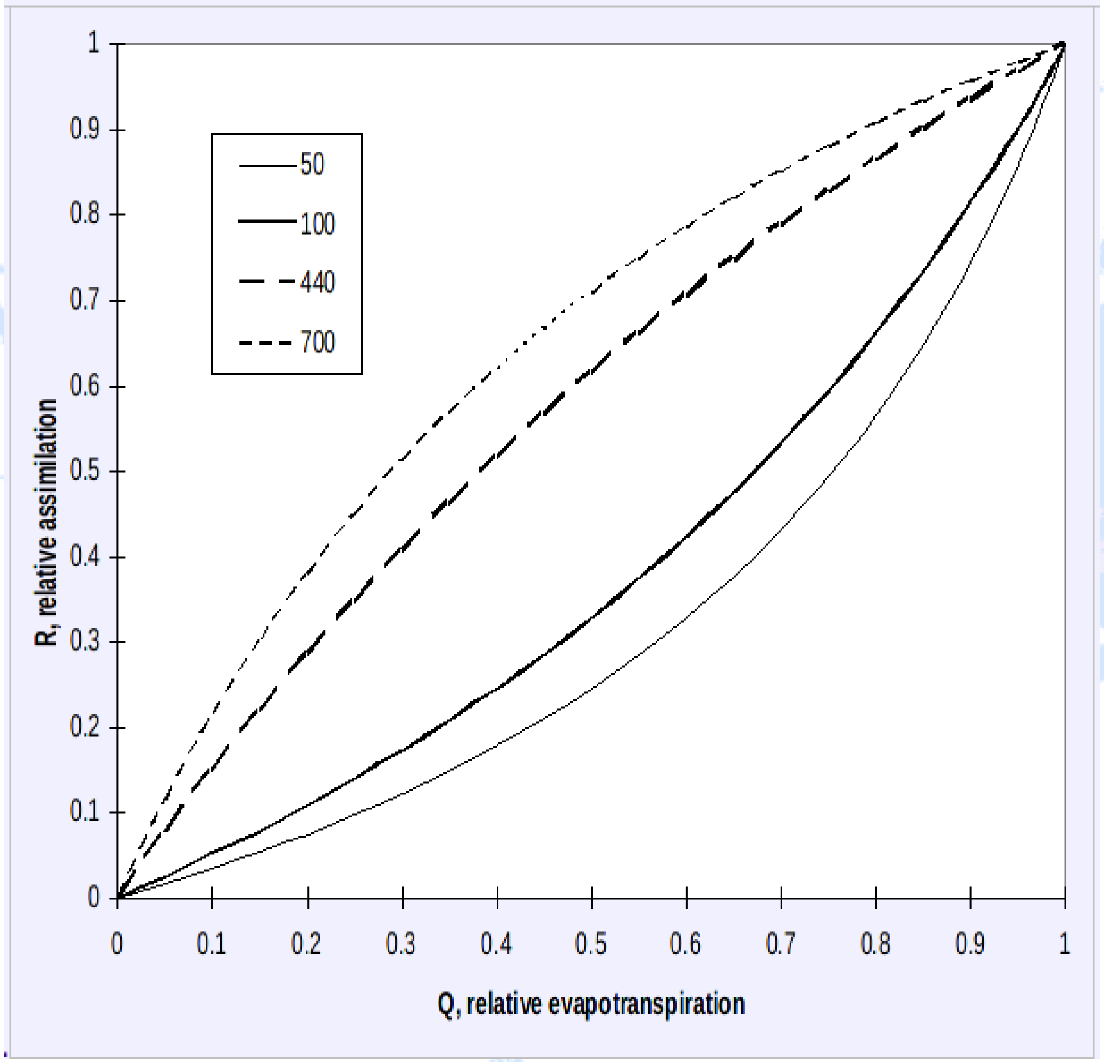
T = ET when LAI>=1

Relative ET as a function of relative soil moisture



Water/CO2 and energy budgets are connected





Water use and photo-synthesis: relative assimilation Vs. relative ET

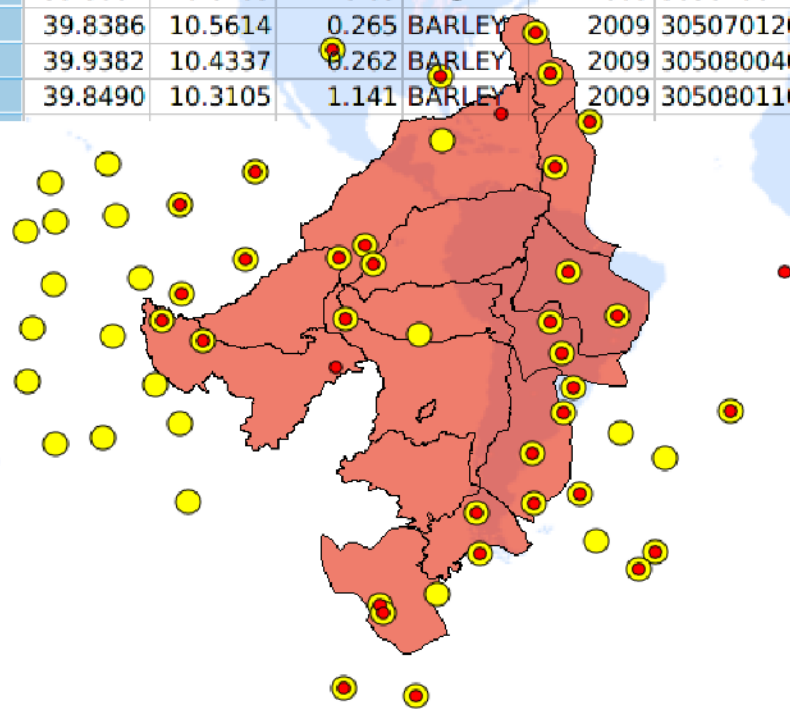
as a function of mesophyll resistance

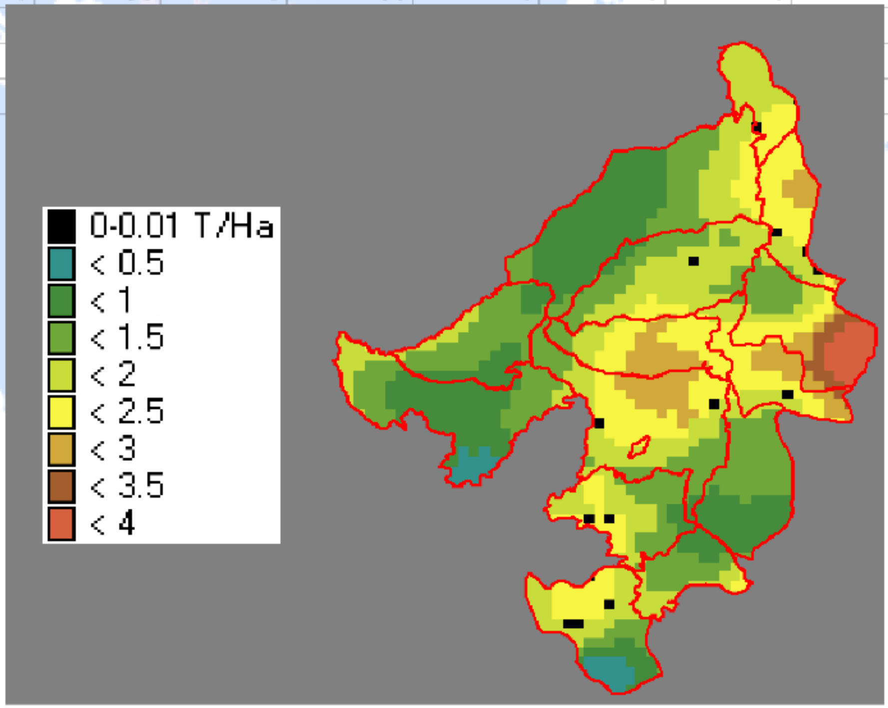
Models, models models, models and models

- Process-oriented simulation models
- Statistical
- Non-parametric
- FAO AgroMetShell
- “Other”

Start with *calibration matrix*

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	lon	lat	yield	crop	year	RK_EA	pc_impr	pc_nat	pc_chem	pc_pure	pc_tmr	pc_tmr_l	pc_tlr	pc_tlr_los
196	39.4017	10.9162	1.725	BARLEY	2009	3040902502	0	0	0	100	0	0	0	0
197	39.1737	10.6998	0.555	BARLEY	2009	3041300603	0	28	0	100	0	0	0	0
198	39.3617	10.5569	2.252	BARLEY	2009	3041400502	12	0	0	100	0	0	0	0
199	39.2460	10.4127	2.26	BARLEY	2009	3041401205	0	0	0	100	0	0	0	0
200	39.5535	10.8537	2.034	BARLEY	2009	3041500704	0	95	0	100	0	0	0	0
201	39.4298	10.5723	2.477	BARLEY	2009	3041501803	0	0	0	100	0	0	0	0
202	39.2787	10.8193	1.783	BARLEY	2009	3042000101	0	31	0	100	2	50	2	25
203													0	0
204													0	0
205													0	0
206													0	0
207													0	0
208	39.2976	10.0710	0.655	BARLEY	2009	3050400502	0	25	25	100	0	0	0	0
209	39.7121	10.4542	1.164	BARLEY	2009	3050500207	0	30	9	100	0	0	4	30
210	39.5607	10.3835	0.673	BARLEY	2009	3050500905	0	37	18	96	0	0	0	0
211	39.5575	10.5525	0.367	BARLEY	2009	3050600603	0	29	0	100	0	0	0	0
212	39.6629	10.5645	1.114	BARLEY	2009	3050600903	0	52	0	100	7	46	0	0
213	39.8002	10.6705	0.69	BARLEY	2009	3050700403	0	33	0	88	0	0	0	0
214	39.8386	10.5614	0.265	BARLEY	2009	3050701204								0
215	39.9382	10.4337	0.262	BARLEY	2009	3050800404								0
216	39.8490	10.3105	1.141	BARLEY	2009	3050801105								0





Regression “models”

Wheat yield (T/Ha) =

$$15.44 + 0.0231\ X_1 - 0.0493\ X_2 + 3.75\ X_4$$

X_1 is November and December rainfall (mm)

X_2 is July average temperature in C

X_4 is July NDVI

Regression "models"

- Calculations are simple and data requirements, limited
- poor performance outside the range of calibration values
- different equations **sometimes** needed for each forecasting time and frequent annual re-calibration

Regression “models”

Best results are achieved with ...

- trend correction
- value added-variables (ETA)
- uncorrelated variables (PCA, fertilizer)
- a variable for maximum local yield (potential yield or other)
- good agronomic analyses better than statistical significance
- “real factors” (NDVI)

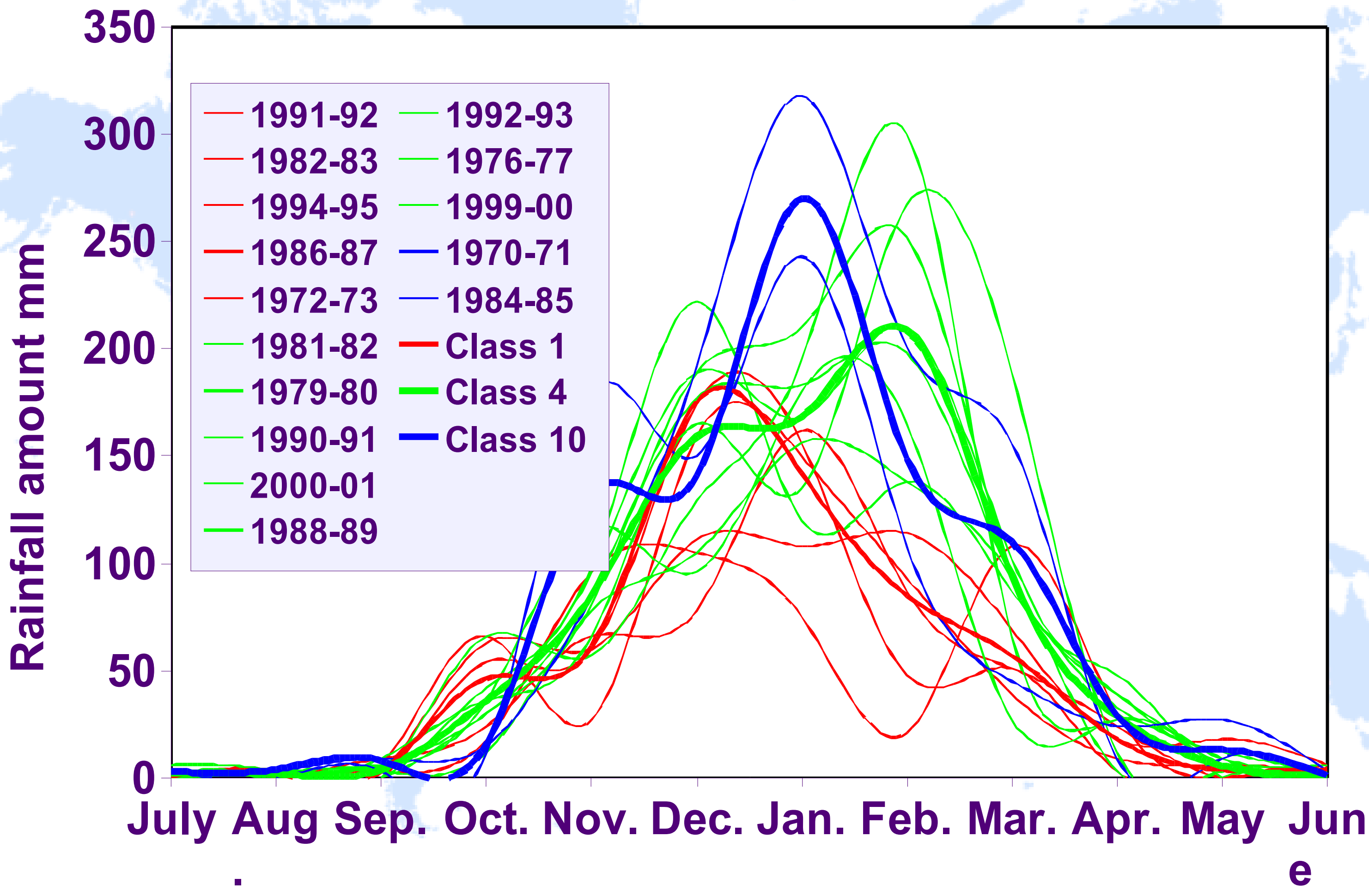
Models, models models, models and models

- Process-oriented simulation models
- Statistical
- Non-parametric
- FAO AgroMetShell
- “Other”

Descriptive: yield T/Ha

March total rainfall	June average sunshine hours per day: 6 hours and less	June average sunshine hours per day: more than 6 hours
75 mm and less	5 ±1	6±2
More than 75 mm	8±1	10±2

Some rainfall profiles (Zimbabwe)



Comparison of methods

Method	R^2		
	Trend	Method	Total
Average Rainfall	0.1702 +	0.4563	0.6265
Water Balance		0.5653	0.7355
Threshold		0.5311	0.7013
Clustering		0.5692	0.7394

Descriptive methods: advantages

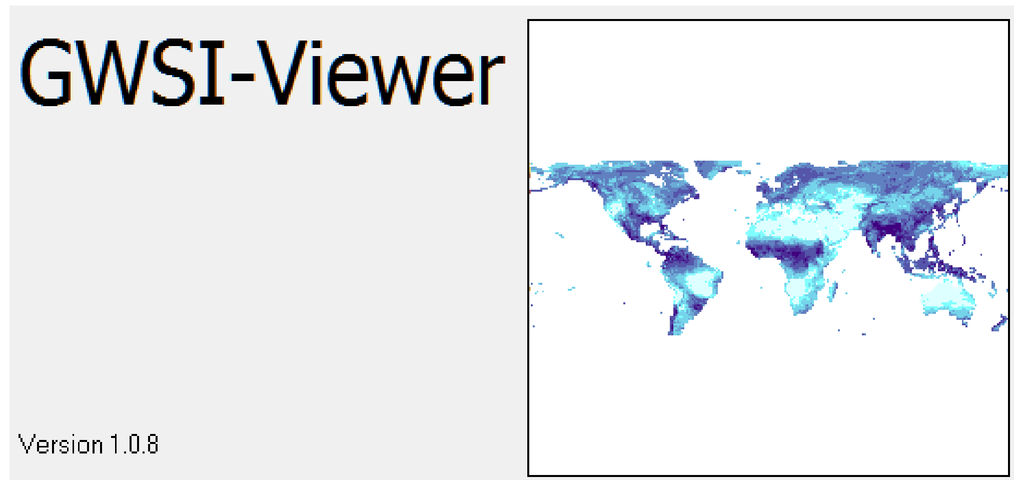
- clustering of combination of mix of time-series and cross-sectional data
- independent of type of functional relation between variables and yield (non-parametric)
- confidence intervals are easy to derive
- require little data processing in operational mode

Models, models models, models and models

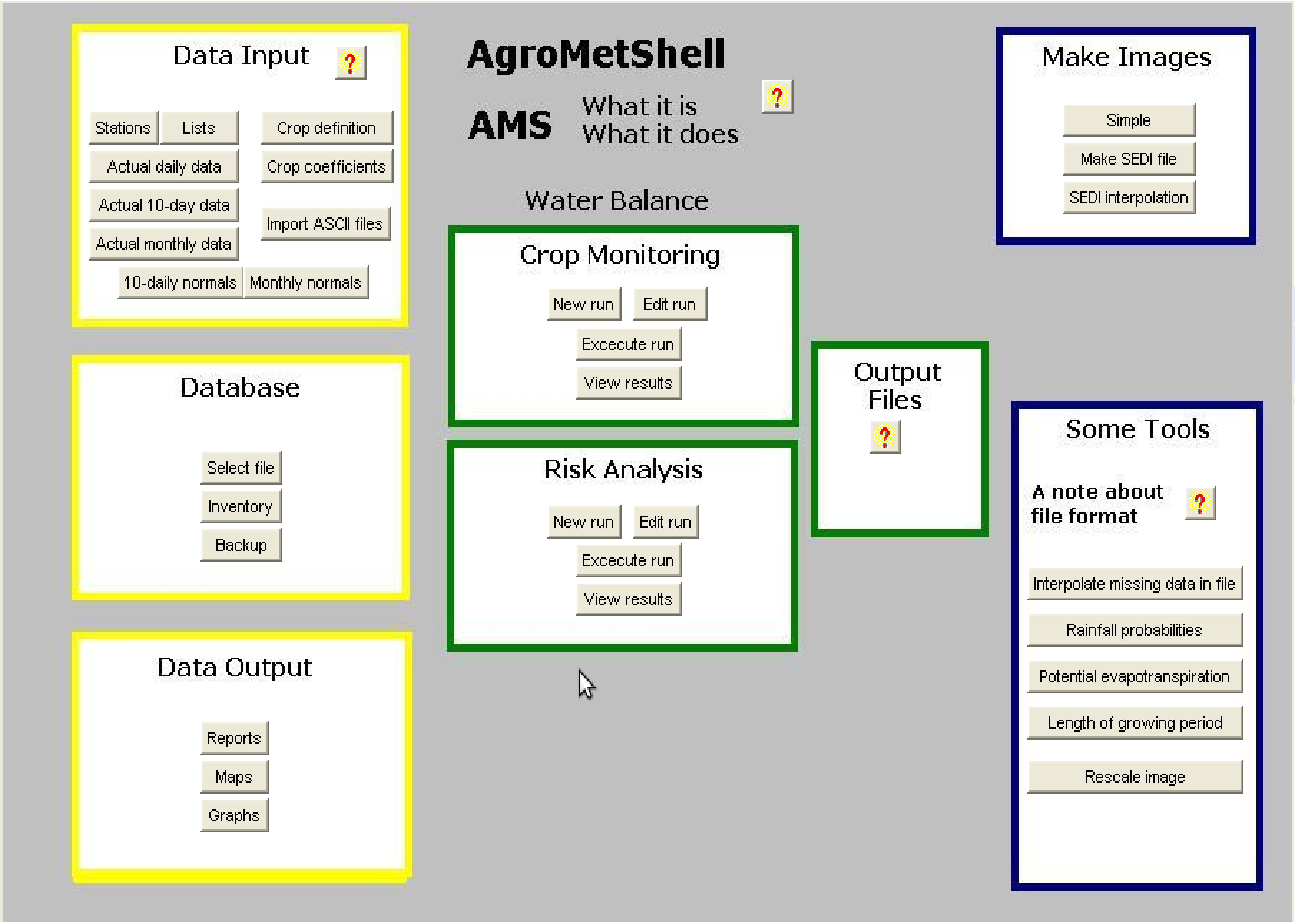
- Process-oriented simulation models
- Statistical
- Non-parametric
- FAO AgroMetShell
- “Other”

wabal

(wb1, wb2, wb3)



wabal (**wa**ter **bal**ance) is a linux/windows command-line impleme
FAO water balance as currently computed in AgroMetShell (FAO),
CMBox[1] (FAO and EC) and some other tools. It is proposed here ir
wb2, wb3) that differ only in the way in which inputs and outputs a


Data Input
Stations
Lists
Crop definition
Actual daily data
Crop coefficients
Actual 10-day data
Import ASCII files
Actual monthly data
10-daily normals
Monthly normals
AgroMetShell
AMS
What it is
What it does
Water Balance
Crop Monitoring
New run
Edit run
Excecute run
View results
Make Images
Simple
Make SEDI file
SEDI interpolation
Database
Select file
Inventory
Backup
Output
Files
Risk Analysis
New run
Edit run
Excecute run
View results
Some Tools
A note about
file format
Interpolate missing data in file
Rainfall probabilities
Potential evapotranspiration
Length of growing period
Rescale image
Data Output
Reports
Maps
Graphs

Overall philosophy of AMS

- Semi-quantitatively assess weather factors relevant for crop production and express them as value-added agronomically meaningful indices (water balance variables, WBV)
- Regress yields against WBVs, and use empirical regression equation for simulation
- A detailed study was done in Morocco using the approach, without major difficulties

Models, models models, models and models

- Process-oriented simulation models
- Statistical
- Non-parametric
- FAO AgroMetShell
- “Other”

"Other" models (in random order!)

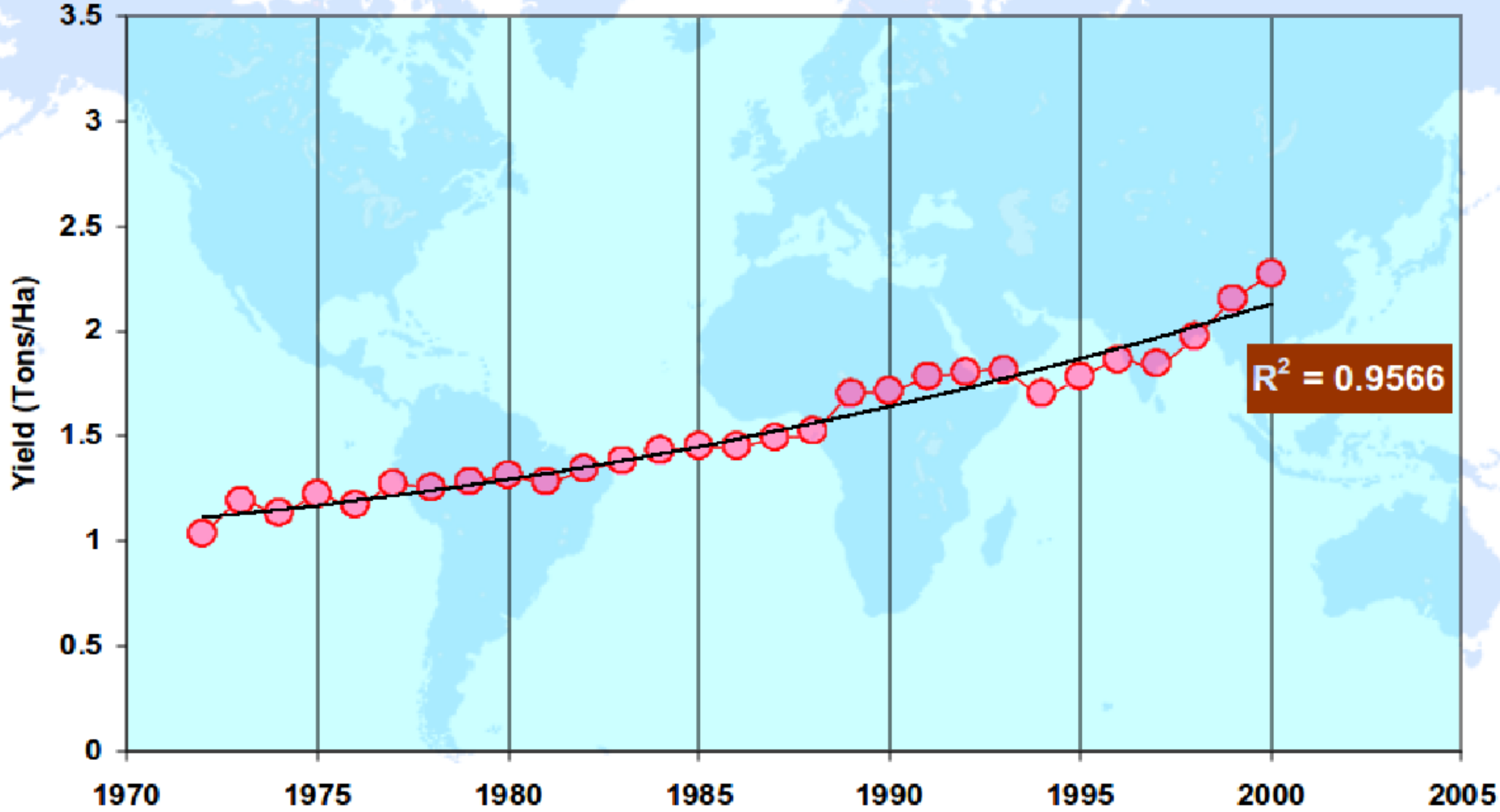
- Subjective with/without a "system"
- Pollen counts
- Counting bales
- CO_2 gradients
- Biometric methods (stem diameter, light interception)

Overview

- There are models, models, models, models and models
- Turning a bit philosophical: some issues, considerations & conclusions

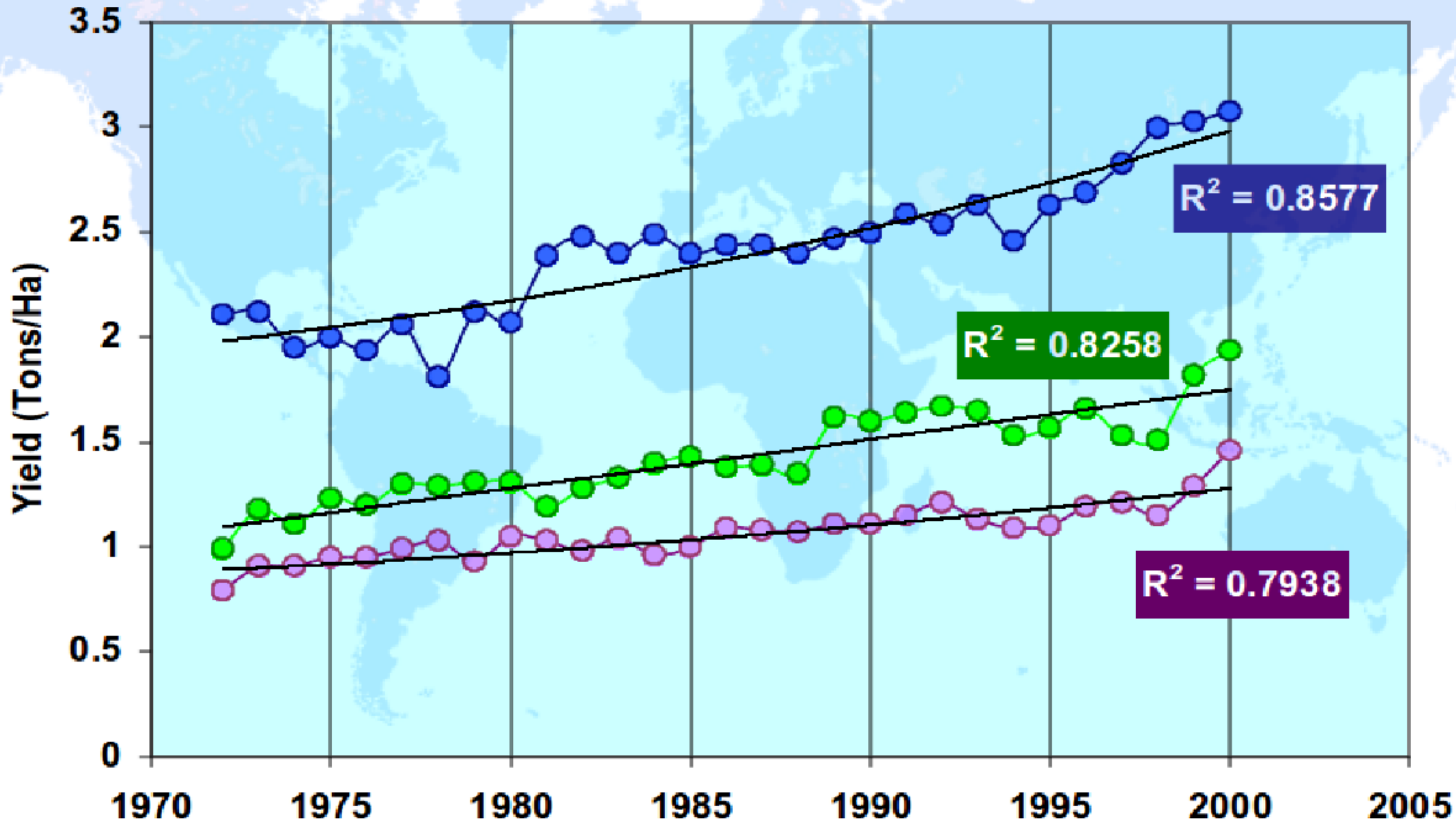
A tool for each scale: national

Total rice yield $R^2=0.96$



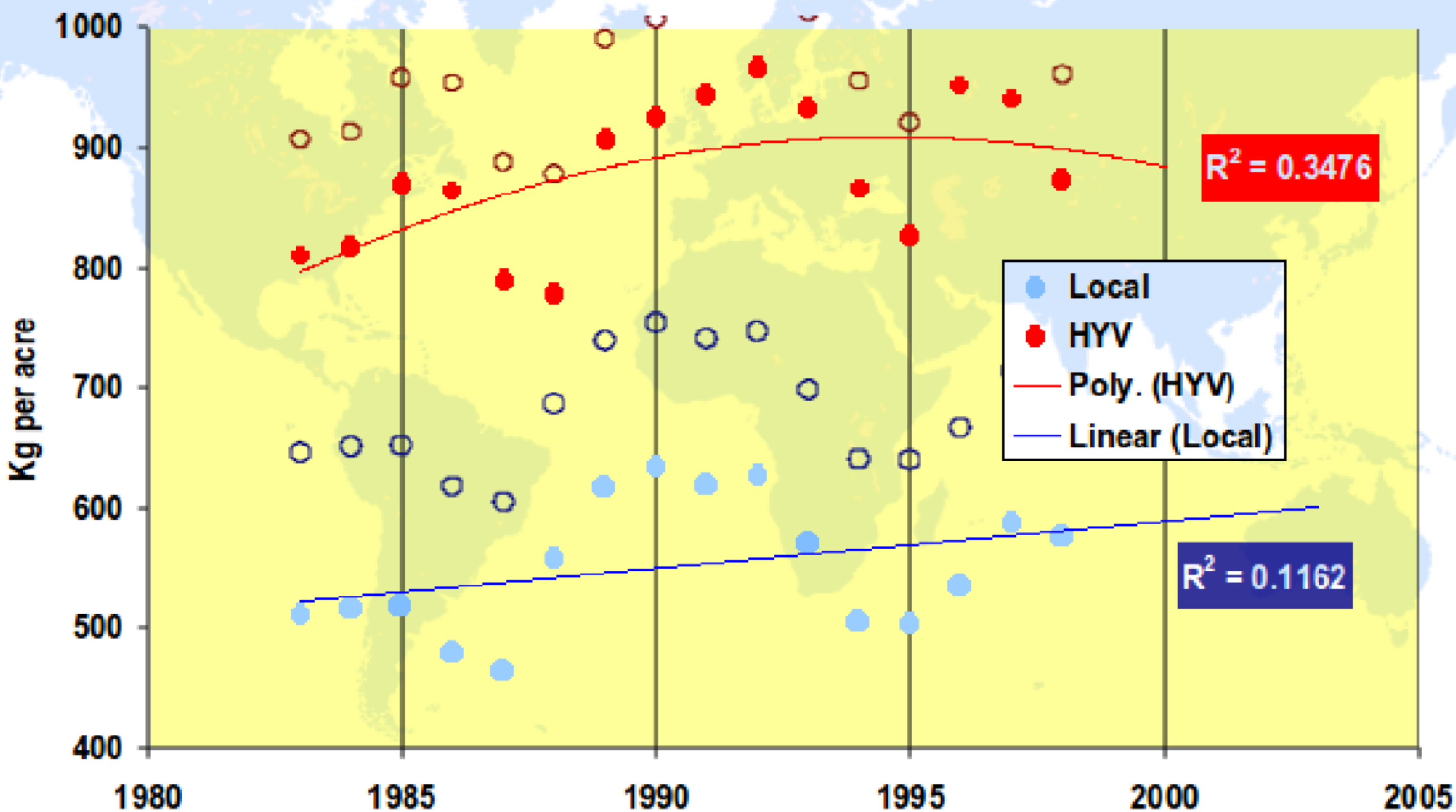
A tool for each scale: typology

Boro/Aman/Aus yield
R^2=0.79 - 0.86



A tool for each scale: local

Rajshahi T-Aman hybrid/local
$R^2 = 0.11 - 0.35$



12 sources of errors: 1-5

1 observation errors in the primary input data

2 processing errors in the input data, including transmission and transcription

3 biases introduced by processing : estimation of missing data, derivation of indirect measures (radar rainfall, radiation...).

4 space and time "scale" errors

5 errors in eco-physiological crop parameters

12 sources of errors: 6 to 8

6 simulation model errors

7 errors due to non-simulated factors (pests, weather at harvest)

8 errors in the agricultural statistics used for the calibration

12 sources of errors: 9 to end

9 calibration errors (choice of statistical relation between crop model output and agricultural statistics)

10 statistical errors in the “future data”

11“second order” errors (when management decisions use early crop forecasts)

12 conflicts between results of different forecasting techniques

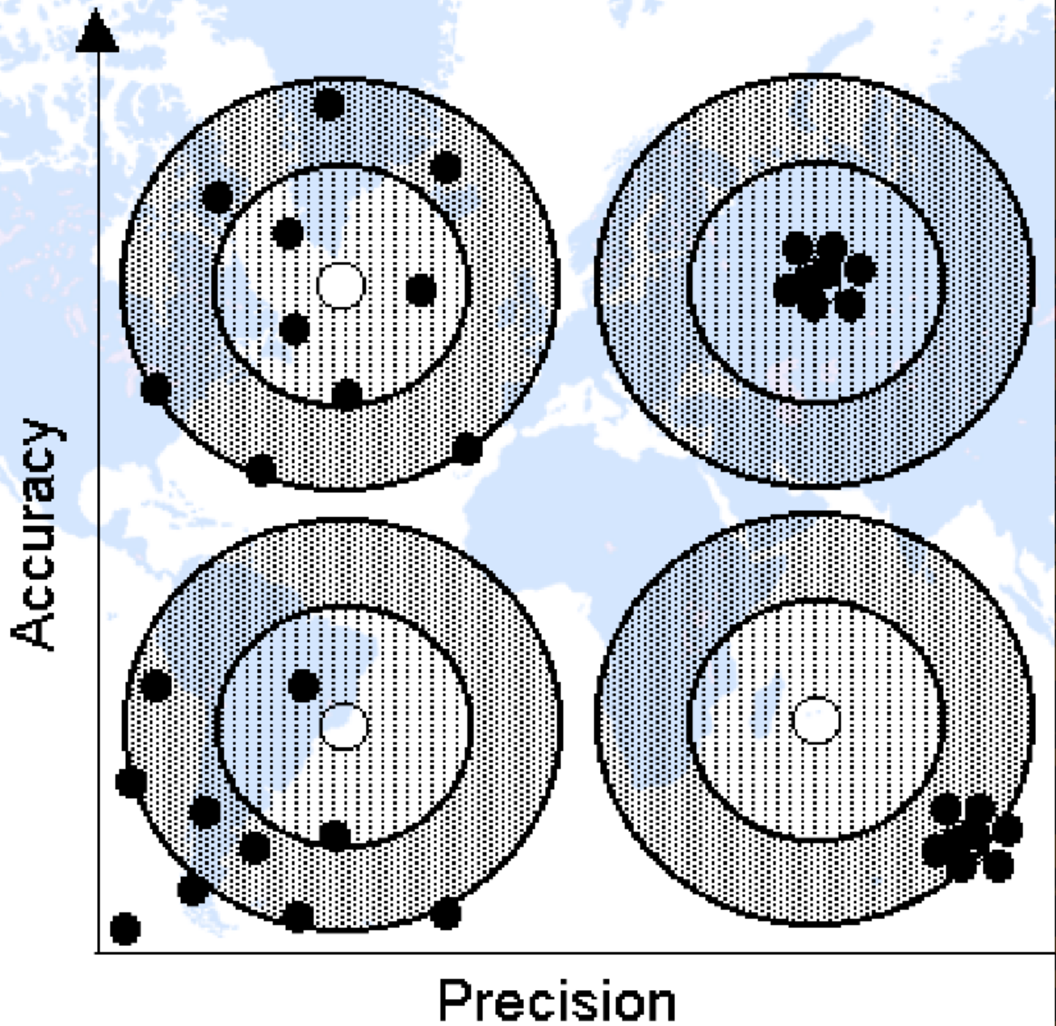
How good is my mo*d*el ? (1/2)

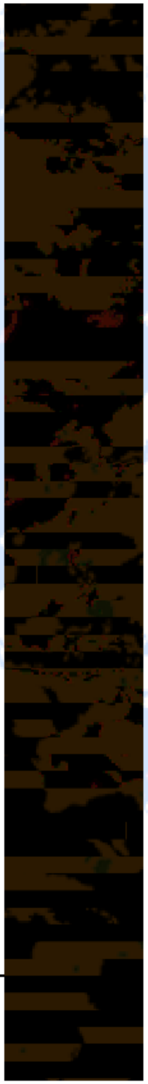
- The standard wording usually resorted to includes *calibration*, *evaluation, validation, verification*
- *Calibration:* adjustment of parameters until the desired results are achieved (proxies!)
- *Evaluation* is descriptive: how realistic, detailed, “balanced/coherent”, “honest” is the model?

How good is my model ? (2/2)

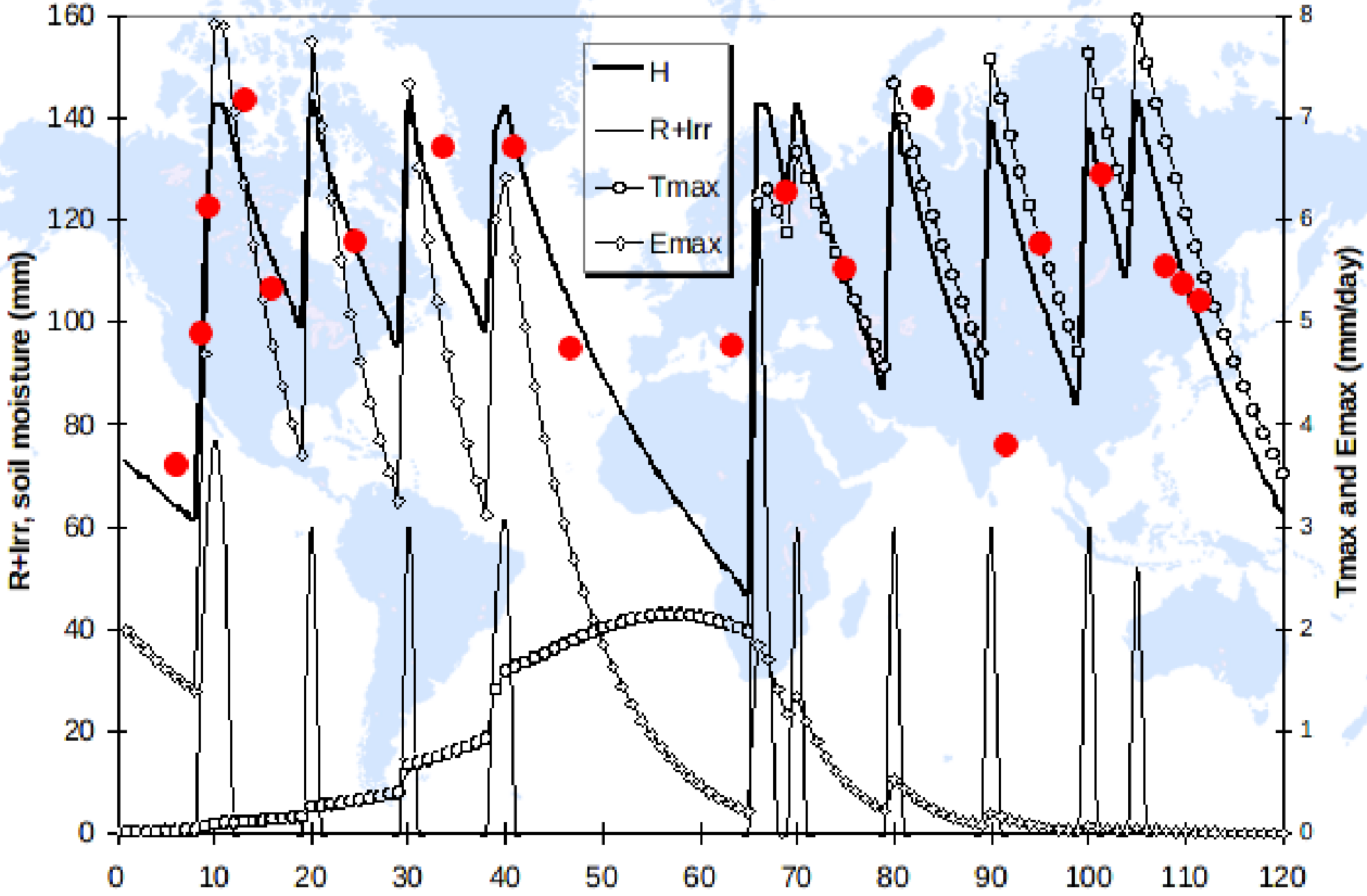
- *Validation*: a model must be validated at the same spatial scale and with the same type of data as those that will be available in operational work; it is the sequence of tests and checks that convince the user that the model is suitable for the intended purpose
- *Verification* (*post-factum*): after I have been using the model for some time and I trust it, if it is accurate and precise, I consider it is *verified*.

Accuracy and precision

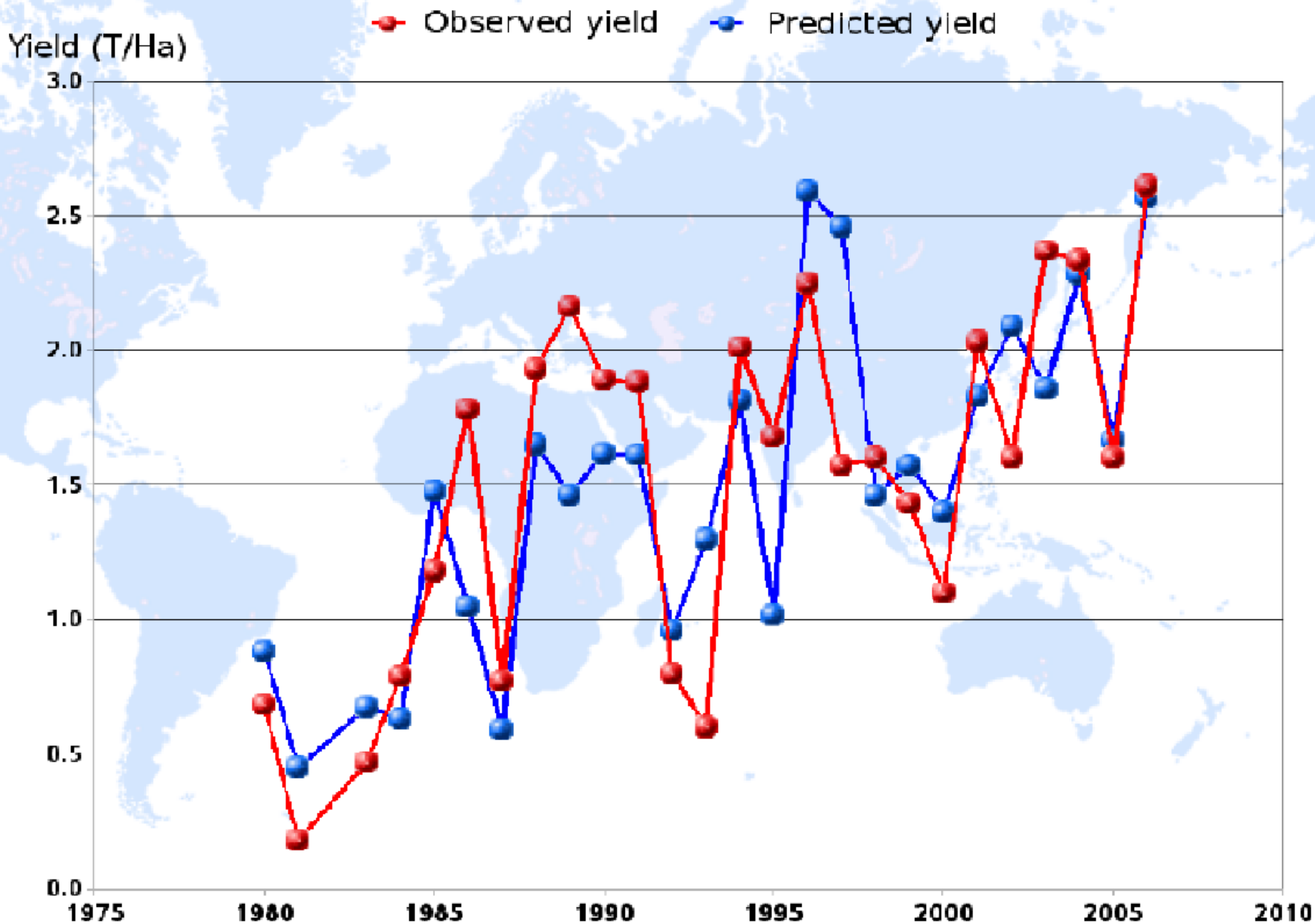




How far can validation be stretched?



How far can validation be stretched? (irrigated *durum*, Morocco)



Conclusions (1/2)

- There is no shortage of simple and complex models to simulate/forecast crop yields for all situations
- Model choice is conditioned/constrained more by data availability than by lack of tools
- Crop modelling – and especially forecasting – is art as much as science

Conclusions (2/2)

- No model remains good forever; in fact, few models remain “good” for more than 3 years!
- Model calibration, validation and verification are mostly subjective exercises
- There is no absolute hierarchy of models; model suitability & quality must be judged based on experience

Thank you!