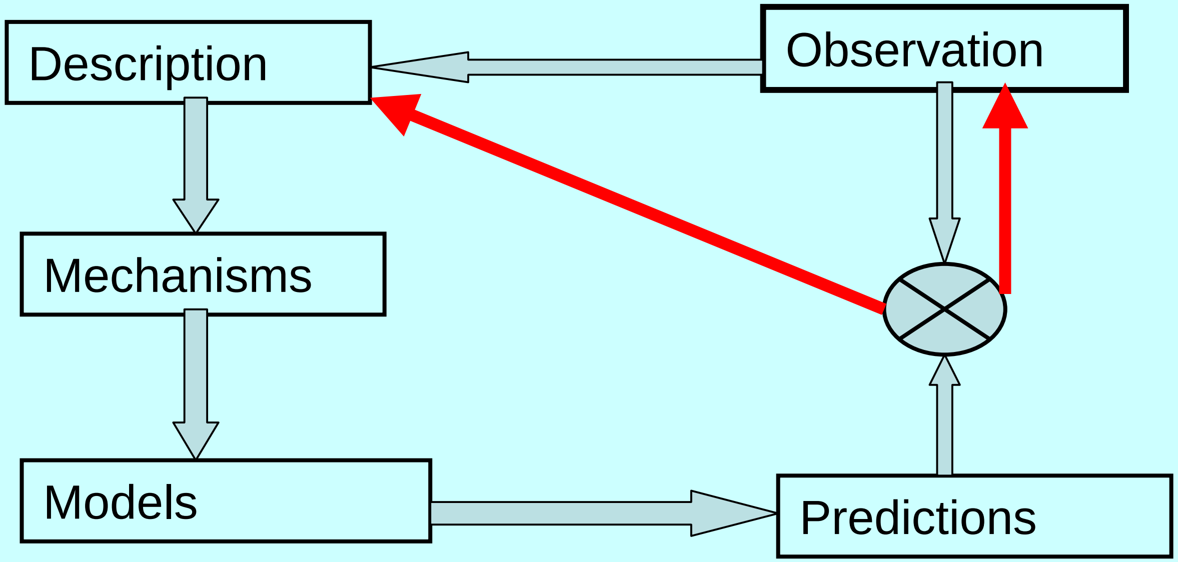
Introduction to CSIRO Biosphere Model (CBM)

Yingping Wang and Ray Leuning
CSIRO Marine and Atmospheric Research, Australia

- C1; time
- C2: incoming short-wave (W/m2)
- C3, net radiation (W/m2)
- C4, air temperature (oC)
- C5, VPD (Pascal), or relative humidity
- C6, windspeed (m/s)
- C7, obs latent heat (W/m2)
- C8, obs sensible heat (W/m2)

Model as a set of hypotheses



Modelling

- Why modelling?
  - Models as a set of hypothesis
  - Models as a synthesis tool
  - Interactions between modelling and measurements

Use of surface flux models for interpreting eddy flux measurements: some basic principles

- Absorbed radiation drives surface processes
- Conservation of mass and energy
- Energy partitioning: demand and supply
- Stomatal functioning

Energy partitioning: the demand and supply

- Energy partitioning: $R_n = \lambda E + H + G$

$$\textbf{\textit{Bowen ratio :}}\ \beta = \frac{H}{\lambda E}$$

β =0.3

β =10

Overview of CBM

- CBM (CSIRO Biosphere Model) simulates exchange of heat, water and CO2 between land surface and atmosphere
- Key processes:
    - Radiative transfer
    - Leaf energy balance
    - Stomatal conductance
    - Leaf photosynthesis model
    - Plant and soil respiration
    - heat, water transfer in soil and snow

two - leaf canopy model

The general structure of CBM

Atmospheric forcings

Canopy radiation;
sunlit & shaded
visible &
near infra-red,
albedo

SEB & fluxes;
for soil-vegetation
system; T_f
λE_f , H_f , λE_g , H_g;
transpiration

stomata transp
& photosynthesis

Plant Carbon
fluxes;
GPP
NPP
NEP

2-leaf canopy

soil temp.

soil moisture

snow

Soil+snow

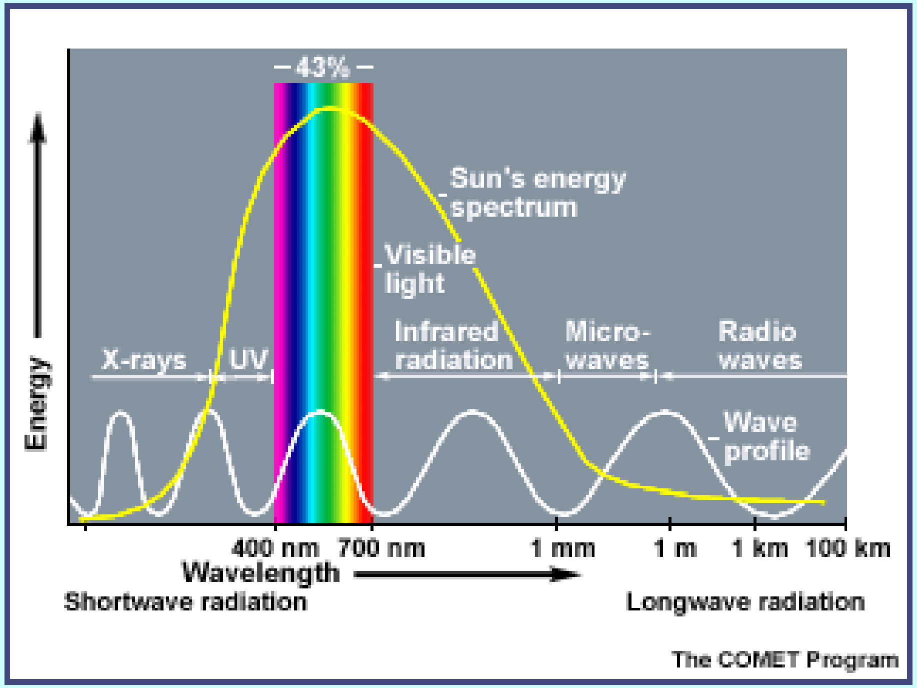
carbon pools; allocation & flow

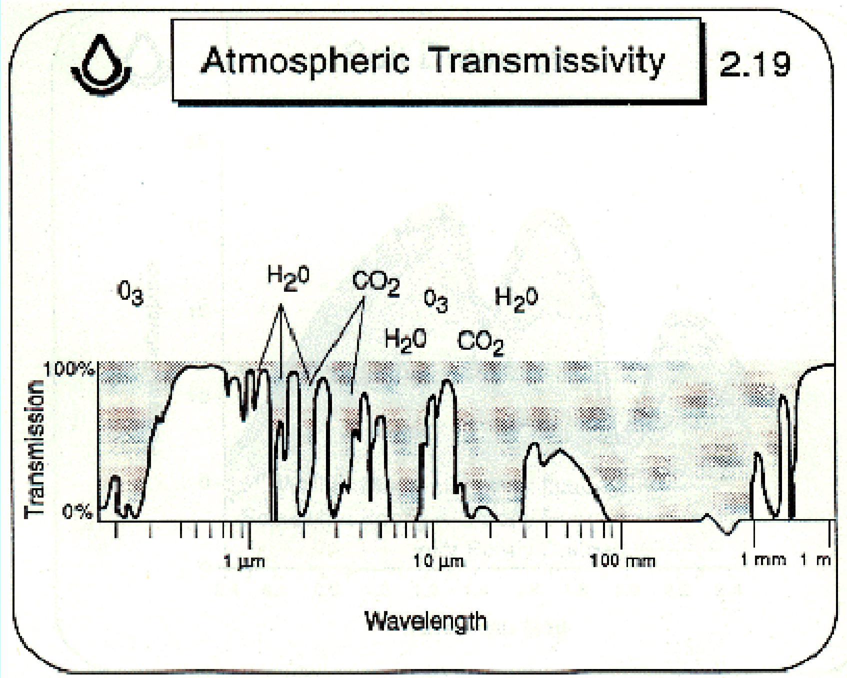
BGC

The two-leaf canopy

- Why two-leaf approach?
  - Multi-layered canopy requires more computing
  - One-leaf approach is inaccurate
- Essence of two-leaf canopy
  - Bulk parameter formulation for sunlit and shaded leaves separately
  - Same equations for single leaf is used for big leaf

Solar radiation and its spectra





Radiation flux density

- The energy unit for radiation is Joule m^{-2} s^{-1}, or Watt m^{-2};
- For photosynthesis, it is not the energy, but number of photos important for the photosystems in a leaf
- The amount of energy per photo decreases with an increase in wavelength. On average

1 W m^{-2} = 4.6 μmol m^{-2} s^{-1} for visible

Four radiation wavebands

- Three radiation wavebands of solar radiation (or shortwave radiation):
- Solar radiation (short-wave radiation)
  - Ultraviolet (0.2 to 0.4 μm); 5-8%
  - Visible (0.4 to 0.7 μm), 46-50%
  - Near infra red (0.7 to 1.5 μm) 44-46%
- Long-wave radiation >10 (μm)

Leaf optical properties

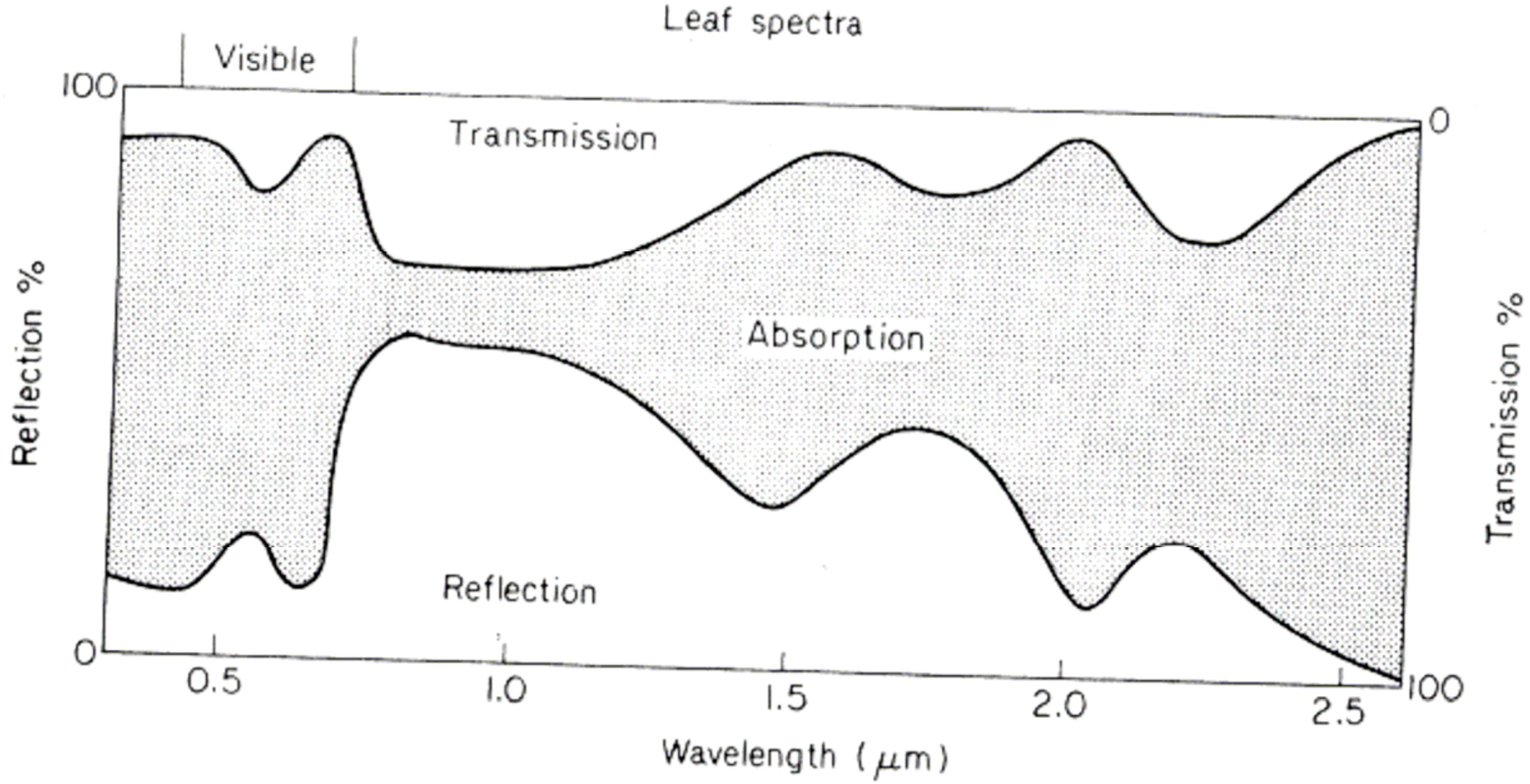


Fig. 6.5 Idealized relation between the reflectivity, transmissivity and absorptivity of a green leaf.

Surface radiation balance

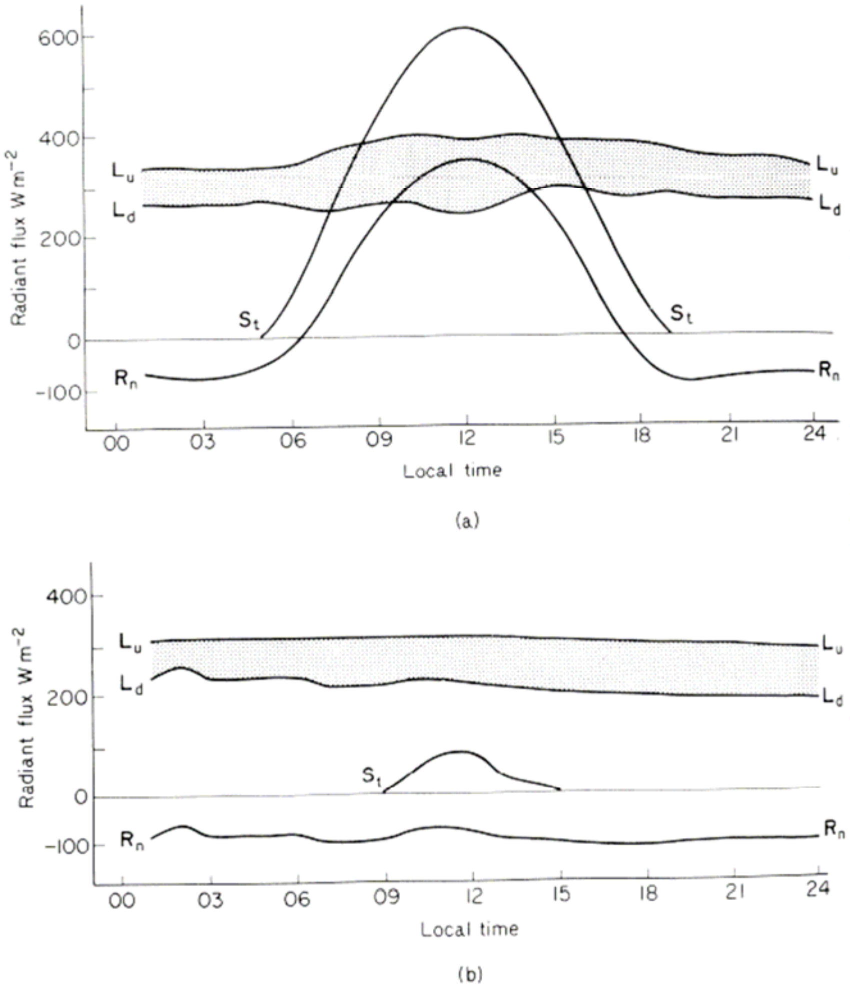
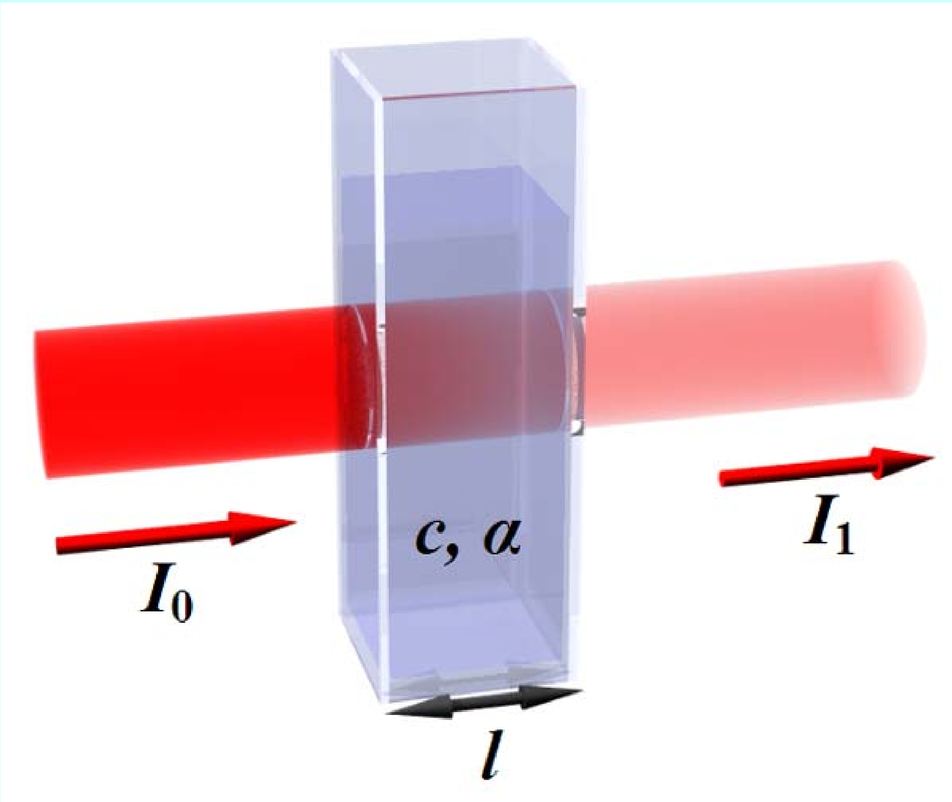


Fig. 4.12 Radiation balance at Bergen, Norway (60°N, 5°E): (a) on 13 April 1968, (b) on 11 January 1968. The grey area shows the net long wave loss and the line $\mathbf{R}_n$ is net radiation. Note that net radiation was calculated from measured fluxes of incoming short and long wave radiation, assuming that the reflectivity of the surface was 0.20 in April (e.g. vegetation) and 0.70 in January (e.g. snow). The radiative temperature of the surface was assumed equal to the measured air temperature.

Beer's law



$$\frac{I_1}{I_0} = \exp(-alc) = \exp(-kL)$$

Where

a: is absorption coefficient;

l: path length;

c: concentration of absorbant;

k: extinction coefficient;

L: canopy leaf area index.

Extinction coefficient (k)

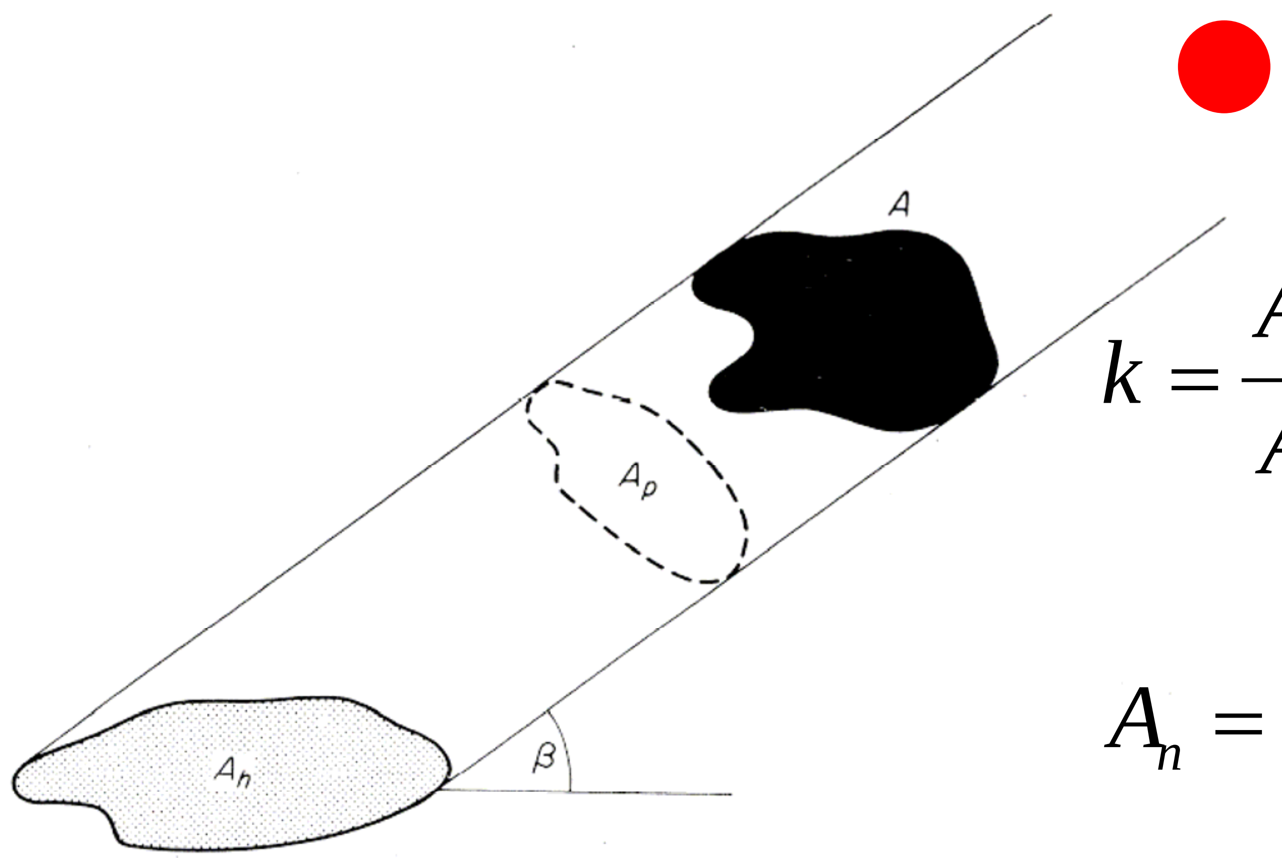


Fig. 5.1 Area A projected on surface at right angles to solar beam (A_p) and on horizontal surface (A_h).

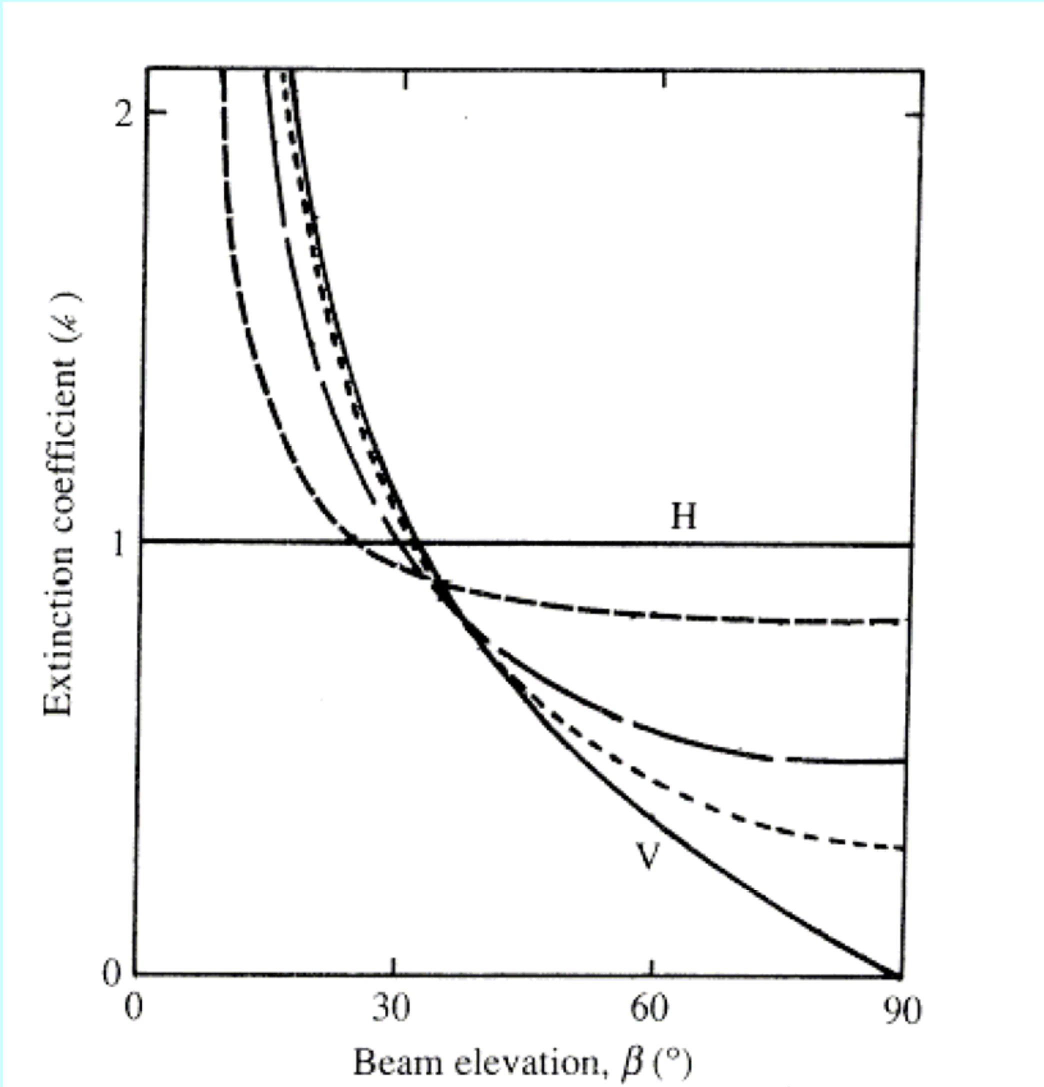
$$k = \frac{A_p}{A_n}$$

$$A_n = A / \sin \beta$$

Leaf angle distribution and k_b

Leaf angle distribution	k_b
Horizontal leaves	$k_b=1$
Vertical leaves	$k_b=2\cot\beta/\pi$
Spherical leaves	$k_b=1/(2\sin\beta)$
Ellipsoidal leaves	$k_b=(x^2+\cot^2\beta)^{0.5}/(A(x)x)$

Extinction coefficient for direct beam radiation (k_b)



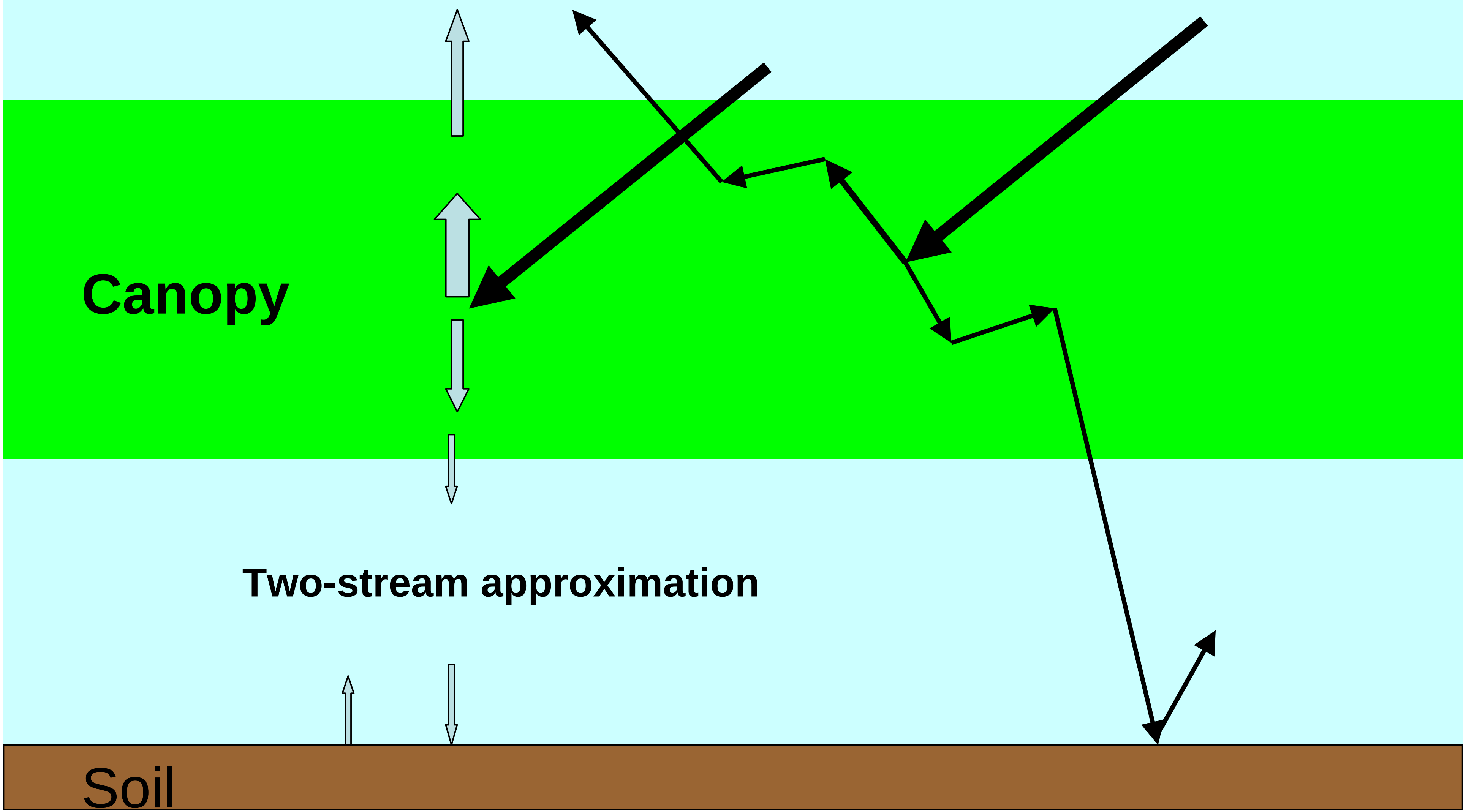
Sunlit leaf area fraction

- Fraction of sunlit leaf area, f_{sun} is equal to gap fraction, $\exp(-k_b L)$.

- For a canopy, total sunlit leaf area index, L_{sun}, is given by

$$L_{sun} = \int_0^L \exp(-k_b \xi) d\xi = \left(1 - \exp(-k_b L\right) / L$$

The two-stream approximation (Goudriaan's model)



Two-stream approximation

Analytic solution to two-stream approximation suggests:

- the flux density of (unintercepted and scattered) diffuse radiation decreases exponentially with the exponent being $k_d^* \xi$, where ξ is cumulative canopy LAI from the canopy top, and

$$k_d^* = k_d \sqrt{1-\omega}.$$

Radiation absorption within a plant canopy

Radiation absorbed by the shaded leaves, q_{shade}

$$q_{shade} = \underbrace{I_d k_d^*(1-\rho_d)\exp(-k_d^*\xi)}_{\textit{Absorbed diffuse radiation}}$$

$$+ \underbrace{I_b\left[k_b^*(1-\rho_b)\exp(-k_b^*\xi) - k_b(1-\omega)\exp(-k_b\xi)\right]}_{\textit{absorbed scattered direct beam radiation}}$$

Radiation absorbed by the sunlit leaves, q_{sun} :

$$q_{sun} = q_{shade} + I_b k_b(1-\omega)$$

Total amount of radiation absorbed

$$\text{All sunlit leaves}, Q_{sun}$$

$$Q_{sun} = \int_0^L \exp(-k_b \xi) q_{sun}(\xi) d\xi \, .$$

$$\text{All shaded leaves}, Q_{shade}$$

$$Q_{shade} = \int_0^L \left(1 - \exp(-k_b \xi)\right) q_{shade}(\xi) d\xi \, .$$

Leaf energy balance

$$R_{lu} + H + \lambda E = R_n$$

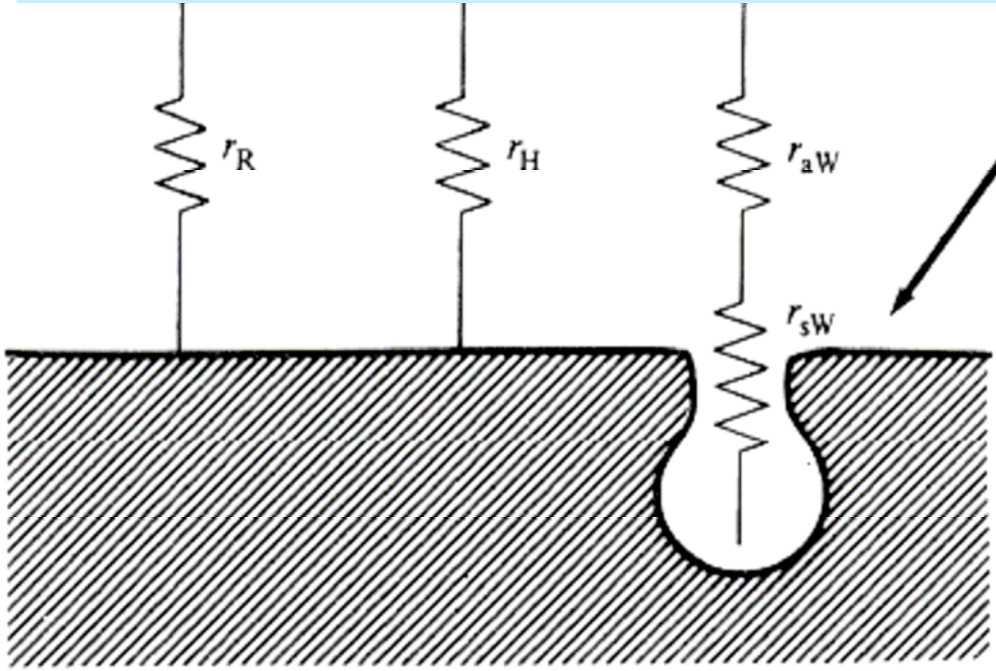


Fig. 5.1. Energy exchanges for a leaf, where the radiative heat loss R_{lu} is the difference between actual net radiation and isothermal net radiation.

Leaf energy balance

- The governing equation:

$$R_{nf} = \lambda E_f + H_f$$

- Sensible heat:

$$H_f = c_p \rho_a (T_f - T_a) g_h$$

- Latent heat (Penman-Monteith equation):

$$\lambda E_f = \frac{s R_{nf,i} + D_a c_p \rho_a g_h}{s + \gamma \frac{g_h}{g_w}}$$

Conductance for heat, water and CO_2

$$g_h^{-1} = g_a^{-1} + (0.5/g_b)^{-1}$$
$$g_w^{-1} = g_s^{-1} + (1.075g_b)^{-1} + g_a^{-1}$$
$$g_c^{-1} = (1.57g_s)^{-1} + (1.27g_b)^{-1} + g_a^{-1}$$

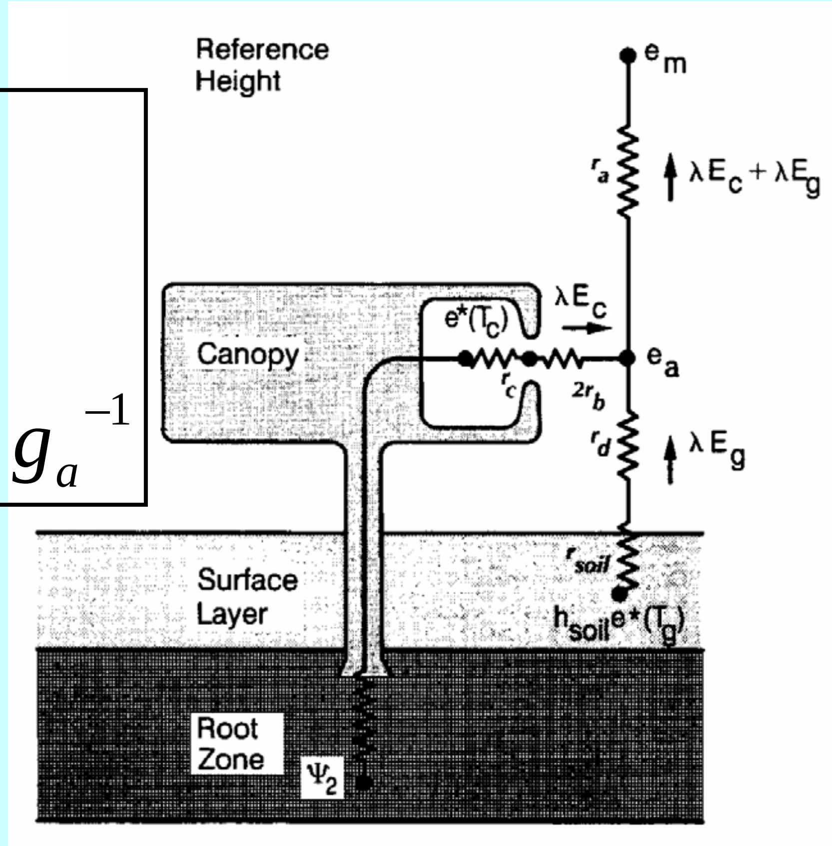


Table 5.1. *Temperature dependence of the 'radiative' conductance g_R, and some typical values of the total thermal conductance (g_{HR}) for a range of values of g_H. The value in brackets is g_H as a percentage of g_{HR}*

		g_{HR} (mm s^{-1})		
Temperature (°C)	g_R (mm s^{-1})	$g_H = 2$	20	200 (mm s^{-1})
0	3.54	5.5 (36)	23.5 (85)	204 (98)
10	4.10	6.1 (33)	24.1 (83)	204 (98)
20	4.69	6.7 (30)	24.7 (81)	205 (98)
30	5.37	7.4 (27)	25.4 (79)	205 (97)
40	6.10	8.1 (25)	26.1 (77)	206 (97)

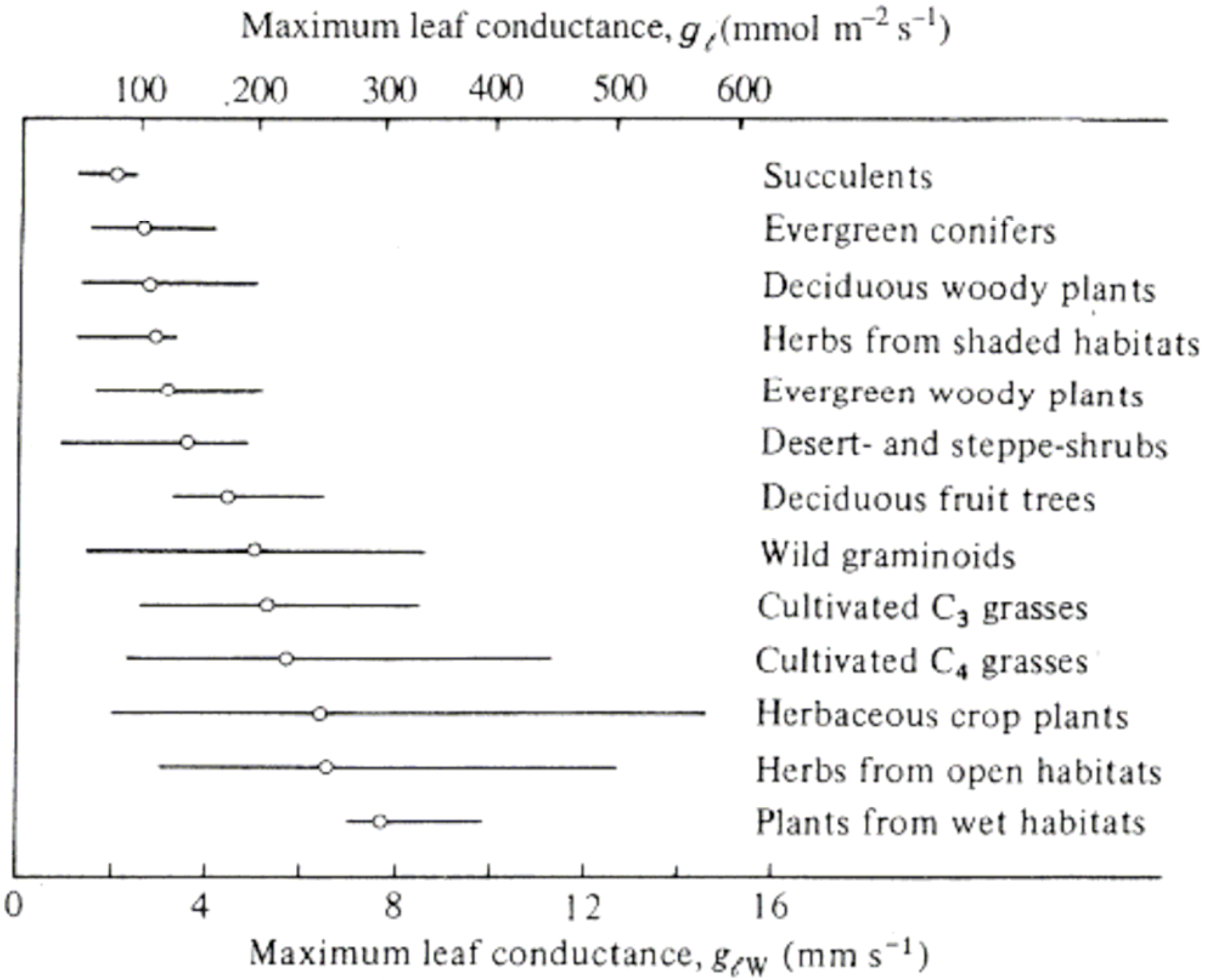


Fig. 6.6. Maximum leaf conductance ($g_{\ell W}$) in different groups of plants. The lines cover about 90% of individual values reported. The open circles represent group average conductances. (Adapted from Körner *et al.* 1979).

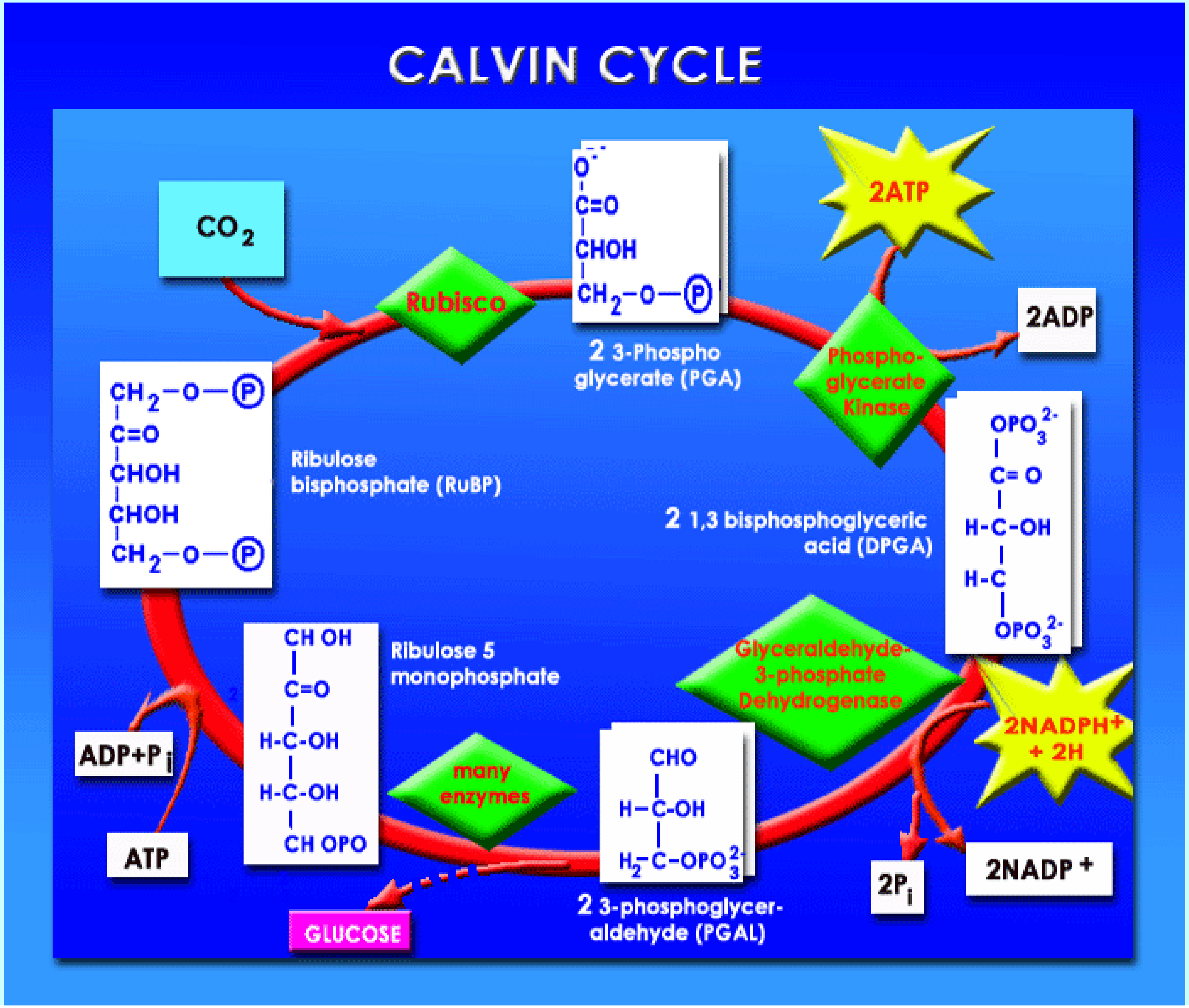
Stomatal conductance: coupling water use and CO_2 uptake

- Through stomata, CO_2 enters and H_2O exits the leaf;
  - When [CO_2] in intercellular space and guard cell ↓ ,K^+ moves into guard cell, stomata opens, vice versa;
  - Too much water loss, stomata close;
  - Soil dries, ABA produced at root tips transported to leaf, and induce stomata closure.

The Ball - Berry - Leuning model

$$g_s = g_0 + \frac{a f_w A_n}{(C_s - \Gamma)\left(1 + D_s / D_0\right)}$$

Leaf photosynthesis: the Calvin cycle



C_3 photosynthesis model

$$A_n = \min\left(\underbrace{V_{c,c}}_{\textit{Rubisco limited}}, \underbrace{V_{c,j}}_{RuBP-\textit{limited}}, \underbrace{V_{c,p}}_{\textit{sink}-\textit{limited}} \right) \left(1 - \underbrace{\Gamma^* / C_i}_{\textit{Photorespiration loss}} \right) - \underbrace{R_d}_{\textit{day respiration}}$$

Rubisco-limited

Light-limited

Sink-limited

$$V_{c,c} = \frac{V_{c\max} C_i}{C_i + K_c \left(1 + O_i / K_o \right)}$$

$$V_{c,j} = \frac{J}{4} \frac{C_i - \Gamma^*}{C_i + 2\Gamma^*}$$

$$V_{c,p} = \frac{V_{c\max}}{2}$$

The combined model of stomatal conductance, photosynthesis and transpiration for a leaf

$$H_f = c_p \rho_a (T_f - T_a) g_h; \qquad \textbf{\textit{unknowns}}: \; T_f \;, H_f$$

$$\lambda E_f = \left(sR_{n,i} + D_a c_p \rho_a g_h \right) / \left(s + \gamma g_h / g_w \right); \; \textbf{\textit{unknows}}: E_f, g_s$$

$$R_n = H_f + \lambda E_f + c_p \rho_a (T_f - T_a) g_r; \qquad \textbf{\textit{unknows}}: H_f, E_f, T_f$$

We have four unknowns with only three equations?

The combined model of stomatal conductance, photosynthesis and transpiration for a leaf

$$A_n = (C_a - C_s) g_{bc}; \qquad \textbf{\textit{unkonws}} : A_n, C_s$$

$$A_n = (C_s - C_i) g_{sc}; \ \textbf{\textit{unknowns}} : A_n, C_s, C_i, g_s$$

$$A_n = V_c(C_i, Q_{PAR}, T_f) - r_d \textbf{\textit{; unknows,}} \ A_n, V_c$$

We have four unknowns, only three equations?

The combined model of stomatal conductance, photosynthesis and transpiration for a leaf

$$g_s = g_0 + \frac{a f_w A_n}{(C_s - \Gamma)\left(1 + D_s / D_0\right)} \quad ; \qquad \textbf{\textit{unknowns}}: g_s, A_n, C_s, D_s$$

$$D_s = D_a + s(T_f - T_a); \qquad \textbf{\textit{unknows}}: D_s, T_f$$

The extra equation forms a link between energy flux and CO_2 exchange. These equations are the core of the combined model

Scaling conductance of leaf to big leaf

Table 1

Formulation of the parameters for the two big-leaf model[a]

$G_{\mathrm{bf},i} = g_{\mathrm{bf}}(0)L_i$

$G_{\mathrm{bu},1} = g_{\mathrm{bu}}(0)\Psi\{0.5k_{\mathrm{u}} + k_{\mathrm{b}}\}$

$G_{\mathrm{bu},2} = g_{\mathrm{bu}}(0)[\Psi\{0.5k_{\mathrm{u}}\} - \Psi\{0.5k_{\mathrm{u}} + k_{\mathrm{b}}\}]$

$$G_{\mathrm{r},1} = \left[\frac{4\sigma T_{\mathrm{a}}^3 k_{\mathrm{d}}\varepsilon_f}{c_{\mathrm{p}}}\right]\left[\Psi\{k_{\mathrm{b}} + k_{\mathrm{d}}\} + \frac{\exp(-k_{\mathrm{d}}L) - \exp(-k_b L)}{k_b - k_d}\right]$$

$$G_{\mathrm{r},2} = \left[\frac{4\sigma T_{\mathrm{a}}^3 k_{\mathrm{d}}\varepsilon_f}{c_{\mathrm{p}}}\right]\left[2\Psi\{k_{\mathrm{d}}\} - \Psi\{k_b + k_{\mathrm{d}}\} - \frac{\exp(-k_d L) - \exp(-k_b L)}{k_b - k_d}\right]$$

---

[a] The total conductances for CO_2, H_2O and heat, $G_{\mathrm{c},i}$ and $G_{\mathrm{h},i}$ are calculated as

$$G_{\mathrm{c},i}^{-1} = G_{\mathrm{a},i}^{-1} + (b_{\mathrm{bc}}G_{\mathrm{b},i})^{-1} + (b_{\mathrm{sc}}G_{\mathrm{s},i})^{-1}$$

$$G_{\mathrm{w},i}^{-1} = G_{\mathrm{a},i}^{-1} + G_{\mathrm{b},i}^{-1} + G_{\mathrm{s},i}^{-1}$$

$$G_{\mathrm{h},i}^{-1} = G_{\mathrm{a},i}^{-1} + (n\mathrm{b}_{\mathrm{bh}}G_{\mathrm{b},i})^{-1}$$

$$\text{and} \quad G_{b,i} = G_{bu,i} + G_{bf,i}$$

where b_{bc}, b_{sc} and b_{bh} are constants required to convert conductances for water vapour to those for CO_2 and heat and where $n = 1$ for amphistomatous leaves and $n = 2$ for hypostomatous ones. For other parameters of the big leaves, see Appendix C.

The combined model of stomatal conductance, photosynthesis and transpiration for a leaf

$H_f = c_p \rho_a (T_f - T_a) g_h$; **unknowns:** T_f **,** H_f

$\lambda E_f = \left(s R_{n,i} + D_a c_p \rho_a g_h\right) / \left(s + \gamma g_h / g_w\right)$; **unknows:** E_f, g_s

$R_n = H_f + \lambda E_f + c_p \rho_a (T_f - T_a) g_r$; **unknows:** H_f, E_f, T_f

$g_s = g_0 + \dfrac{a f_w A_n}{(C_s - \Gamma)\left(1 + D_s / D_0\right)}$ **;** **unknowns:** g_s, A_n, C_s, D_s

$D_s = D_a + s(T_f - T_a)$; **unknows:** D_s, T_f

$A_n = (C_a - C_s) g_{bc}$; **unkonws:** A_n, C_s

$A_n = (C_s - C_i) g_{sc}$; **unknowns:** A_n, C_s, C_i, g_s

$A_n = V_c(C_i, Q_{PAR}, T_f) - r_d$**;** **unknows,** A_n, V_c

Respiration: plants

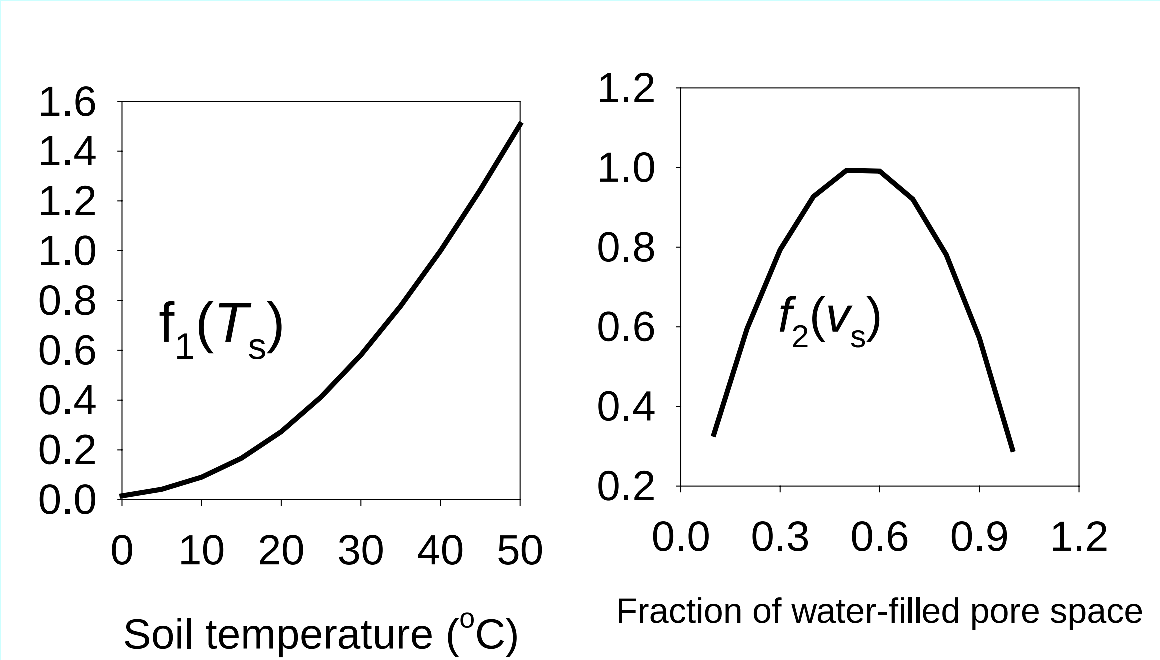
Plant respiration includes growth and maintenance respiration ($R_p = R_g + R_m$)

- Growth respiration (R_g): about 30% of the total carbon for growth is respired;
- Maintenance respiration (R_m): a function of substrate concentration and temperature.
  - $R_m = R_0 \exp(k\mathrm{T})$
  - $k = a - bT$

Respiration: soil

- Soil respiration, R_s, can be modelled as

$$R_s = R_0\, f_1(T_s)\, f_2(v_s)$$



Soil temperature and moisture

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial}\left(\kappa \frac{\partial T}{\partial z}\right) \qquad \textbf{\textit{for temperature}}$$

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial}\left(K \frac{\partial \theta}{\partial z}\right) - \boldsymbol{S}ink + Source \qquad \textbf{\textit{for moisture}}$$

Soil temperature

When thermal conductivity, κ, is constant, and

$$\boldsymbol{T(0,t)} = \boldsymbol{\overline{T}} + A(0)\sin\omega t$$

Solution to the soil temperature equation

$$T(z,t) = \boldsymbol{\overline{T}} + A(0)\exp(-z/D)\sin\left(\omega t - z/D\right)$$

and

$$D = \sqrt{\frac{2\kappa}{\omega}}$$

Soil temperature profile

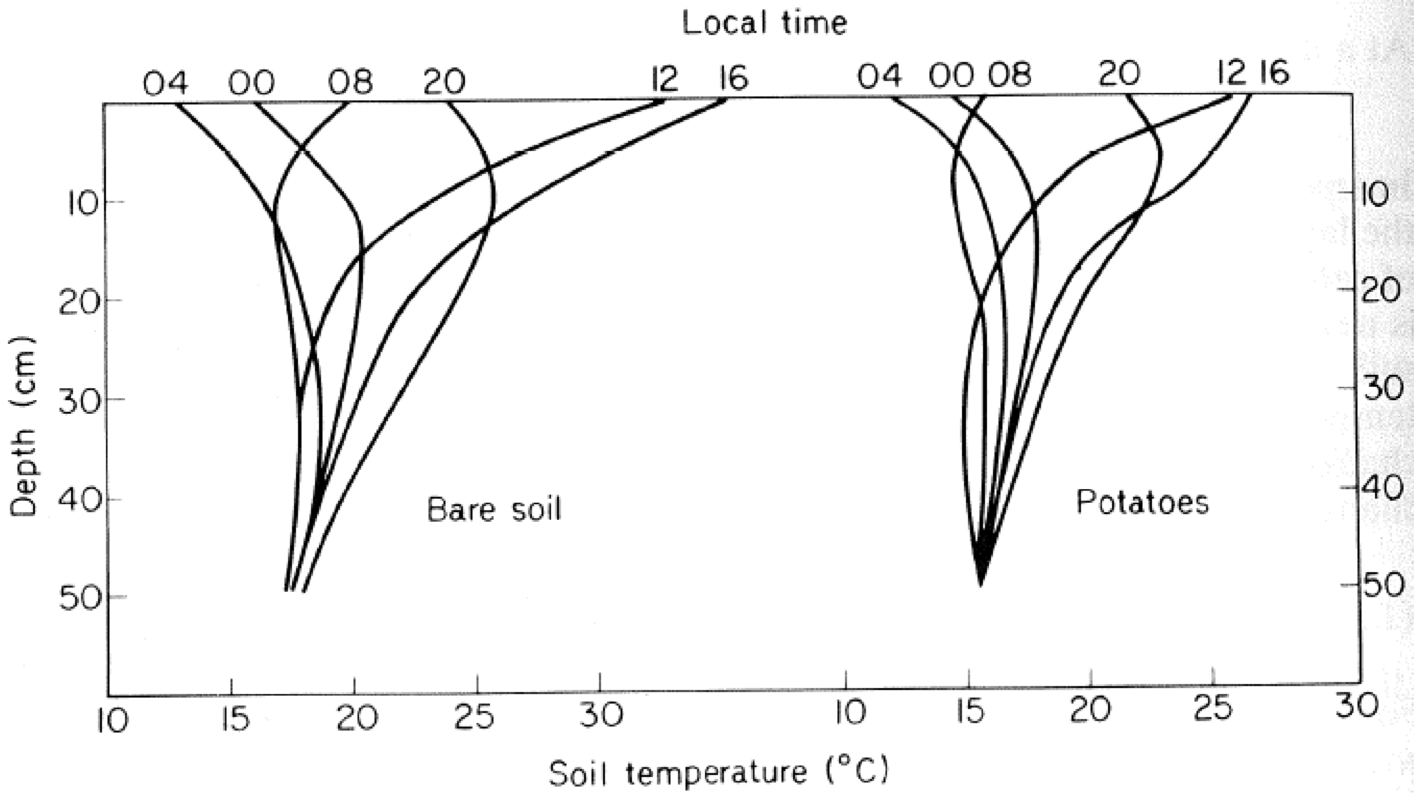
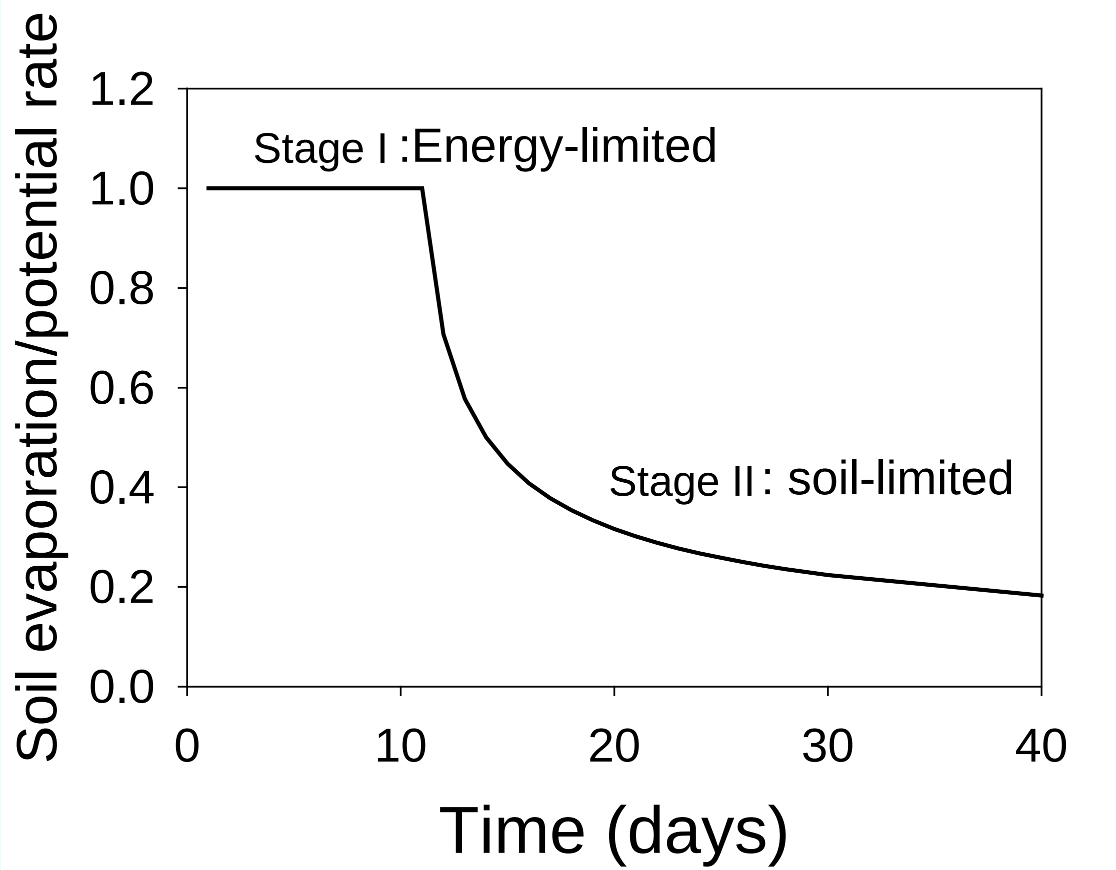


Fig. 13.5 Diurnal change of soil temperature measured below a bare soil surface and below potatoes (from van Eimern, 1964).

Soil evaporation

Two-staged processes



Modelling soil evaporation

$$E_s = \min\left(\underbrace{\rho_a \Delta_q g_a}_{\textit{atmospheric demand}}, \underbrace{\frac{R_{ns} s}{(s+\gamma)\lambda}}_{\textit{energy-limited}}, \underbrace{\frac{\theta_s \Delta h}{\Delta t}}_{\textit{soil supply}} \right)$$

The general structure of CBM

Atmospheric forcings

Canopy radiation;
sunlit & shaded
visible &
near infra-red,
albedo

SEB & fluxes;
for soil-vegetation
system; T_f
λE_f , H_f , λE_g , H_g;
transpiration

stomata transp
& photosynthesis

Plant Carbon
fluxes;
GPP
NPP
NEP

2-leaf canopy

soil temp.

soil moisture

snow

Soil+snow

carbon pools; allocation & flow

BGC

How is it implemented?

Table 2
Structure of the coupled, two-leaf model

Set all physical, physiological constants
Read in location and plant species-dependent parameters
Read in meteorological data (do loop)
Initialise some variables
Calculate radiation absorbed under isothermal conditions
Calculate parameters of the two big leaves
Solve the coupled model (iteration loop for two big leaves)
End of iteration loop
Calculate $A_{c,i}$, $G_{s,i}$, $\lambda E_{c,i}$, $H_{c,i}$
Output results
End do loop

Nonlinear parameter estimation

- An introduction

- Some basic concept
- Linear inversion
- Nonlinear inversion
- Practicals

Inversion

- What is it? You often do it without knowing it.
- Many commercial packages available
- Know your measurements well before inversion
- Often requires a few trials and errors to get the right answer

Some basic concept

- Parameters ($\boldsymbol{p}$), variables ($\boldsymbol{x}$, $\boldsymbol{y}$) and state
- Models ($\boldsymbol{y}=f(\boldsymbol{x},\boldsymbol{p})$)
- Errors: systematic errors and random errors (ε)

Some basic concept

- Maximum likelihood
  - The most probable solution
- Least squares
  - Represent the squared difference, may not be the maximum likelihood solution
- Sensitivity (derivatives)
  - Important for any nonlinear optimization

When is max likelihood solution is the same as least square solution?

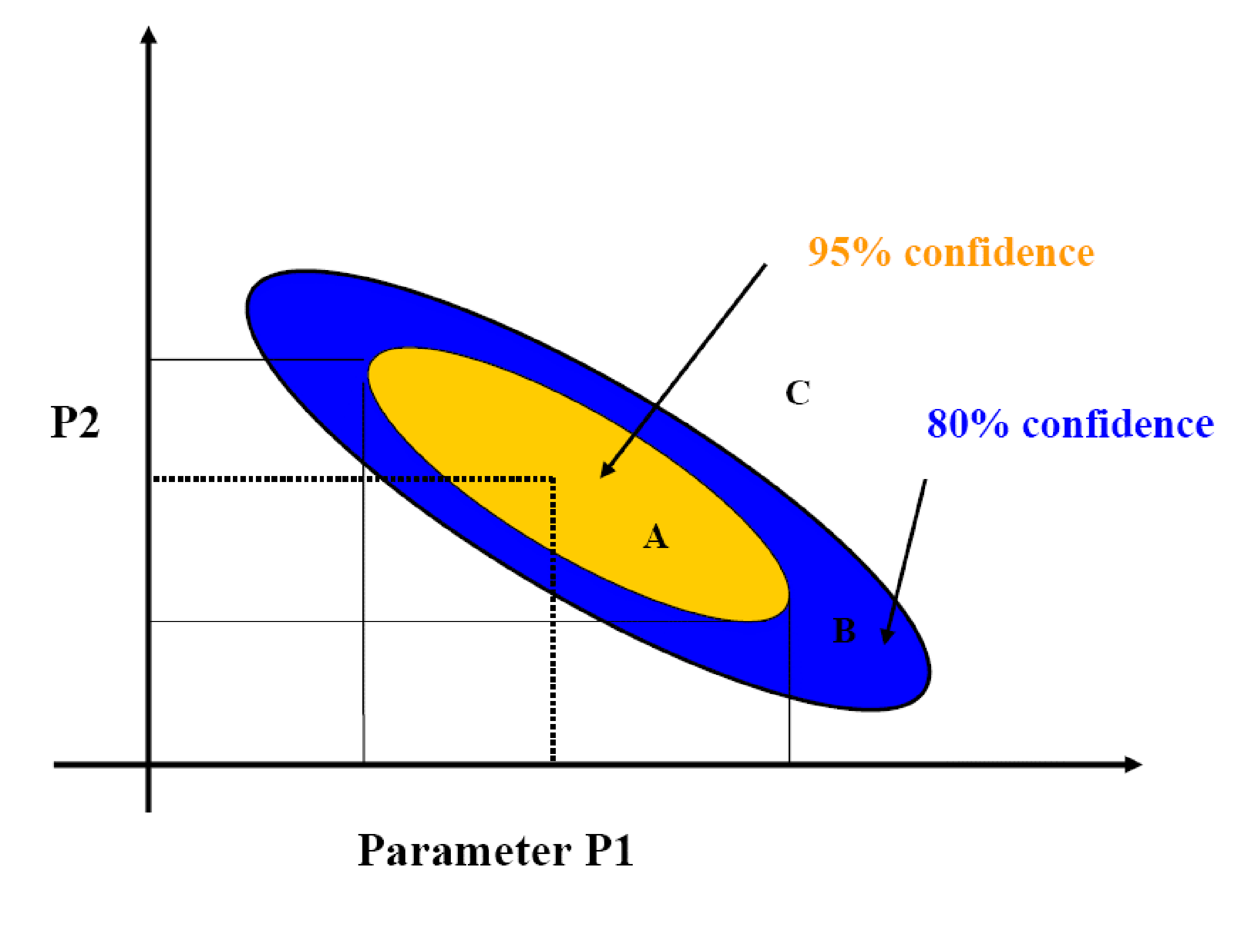
- When the errors of individual data points are normally distributed and independent.

$$P \propto \prod_{i=1}^{N} \left\{ \exp\left[-\frac{1}{2} \left(\frac{y_i - y(x_i)}{\sigma} \right)^2 \right] \Delta y \right\}$$

- Often chi-square is a better distribution for assessing the goodness of fit.

Some basic concepts

- Estimate and probability distribution



Variance and covariance

$$\mathrm{var}(P_1) = \sigma_1^2; \quad \mathrm{var}(P_2) = \sigma_2^2$$

$$\mathrm{var}(P_1 + P_2) = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2$$

Therefore

$$\rho_{12} = 0 \qquad \mathrm{var}(P_1 + P_2) = \mathrm{var}(P_1) + \mathrm{var}(P_2)$$

$$\rho_{12} < 0 \qquad \mathrm{var}(P_1 + P_2) < \mathrm{var}(P_1) + \mathrm{var}(P_2)$$

General linear regression

$$\underbrace{\begin{pmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \vdots \\ \boldsymbol{y_n} \end{pmatrix}}_{\substack{\textit{dependent variable} \\ \textit{n by 1}}} = \underbrace{\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}}_{\substack{\textit{independent variables} \\ \textit{n by m}}} \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\substack{\textit{parameter} \\ \textit{m by 1}}} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}_{\substack{\textit{error} \\ \textit{n by 1}}}$$

In matrix form

$$\boldsymbol{Y} = \hat{Y} + \varepsilon = \boldsymbol{Xb} + \varepsilon$$

Linear inversion theory

- For a given set of measurements of ($\boldsymbol{X}_0$, $\boldsymbol{Y}_0$), the maximum likelihood estimate of coefficient $\boldsymbol{b}$ is given by

$$\boldsymbol{b} = \left(\boldsymbol{X}_0^T \boldsymbol{X}_0\right)^{-1} \boldsymbol{X}_0^T \boldsymbol{Y}_0$$

- The covariance of $\boldsymbol{b}$ (cov($\boldsymbol{b}$)) is given by

$$\mathrm{cov}(\boldsymbol{b}) = \sigma^2 \left(\boldsymbol{X}_0^T \boldsymbol{X}_0\right)^{-1}$$

An example

Y: dependent variable; x_1 and x_2 are two independent variables. The five set of observations are: (x_{1i}, x_{2i}, y_i)

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ 1 & x_{14} & x_{24} \\ 1 & x_{15} & x_{25} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{pmatrix}$$

$$y = \hat{y} + \varepsilon = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon$$

The covariance matrix : $\mathrm{cov}(\boldsymbol{b}) = \sigma^2\left(\boldsymbol{X_0^T X_0}\right)^{-1}$

$$\mathrm{cov}(b) = \begin{pmatrix} \sigma_0{}^2 & \sigma_0\sigma_1 & \sigma_0\sigma_2 \\ \sigma_0\sigma_1 & \sigma_1{}^2 & \sigma_1\sigma_2 \\ \sigma_0\sigma_2 & \sigma_1\sigma_2 & \sigma_2{}^2 \end{pmatrix}$$

$$= \sigma^2\left(\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \end{pmatrix}\begin{pmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ 1 & x_{14} & x_{24} \\ 1 & x_{15} & x_{25} \end{pmatrix}\right)^{-1}$$

Nonlinear inverse theory

- Assume a general nonlinear relationship between ***Y*** and ***X*** with parameter ***p***, and we wish to estimate parameter ***p*** from a set of observations of ($\boldsymbol{X}_0$, $\boldsymbol{Y}_0$).
- The regression model can be written as

$$\boldsymbol{Y} = \hat{\boldsymbol{Y}} + \varepsilon = \boldsymbol{F}(\boldsymbol{X}, \boldsymbol{p}) + \varepsilon$$

Nonlinear inverse theory

- The least squared cost, Φ, is given

$$\phi = \sum_{m} \left(\boldsymbol{Y}_{obs} - \boldsymbol{F}\left(\boldsymbol{X}, \boldsymbol{p}\right)\right)\boldsymbol{Q}\left(\boldsymbol{Y}_{obs} - \boldsymbol{F}\left(\boldsymbol{X}, \boldsymbol{p}\right)\right)^{T}$$

- The optimum is found when

$$\partial \phi \big/ \partial \boldsymbol{p} = 0$$

Nonlinear inverse theory

Using the least square theory, the estimate of parameter ***p***, ***p***$_{es}$, can be calculated as

$$\boldsymbol{p}_{es} = (\boldsymbol{J}^T\boldsymbol{J})^{-1}\boldsymbol{J}^T\left(\boldsymbol{Y}_{obs} - \hat{\boldsymbol{Y}}\right)$$

and the covariance of ***p*** is given by

$$\mathrm{cov}(\boldsymbol{p}_{es}) = \sigma^2(\boldsymbol{J}^T\boldsymbol{J})^{-1}$$

What does it mean?

- Linear:

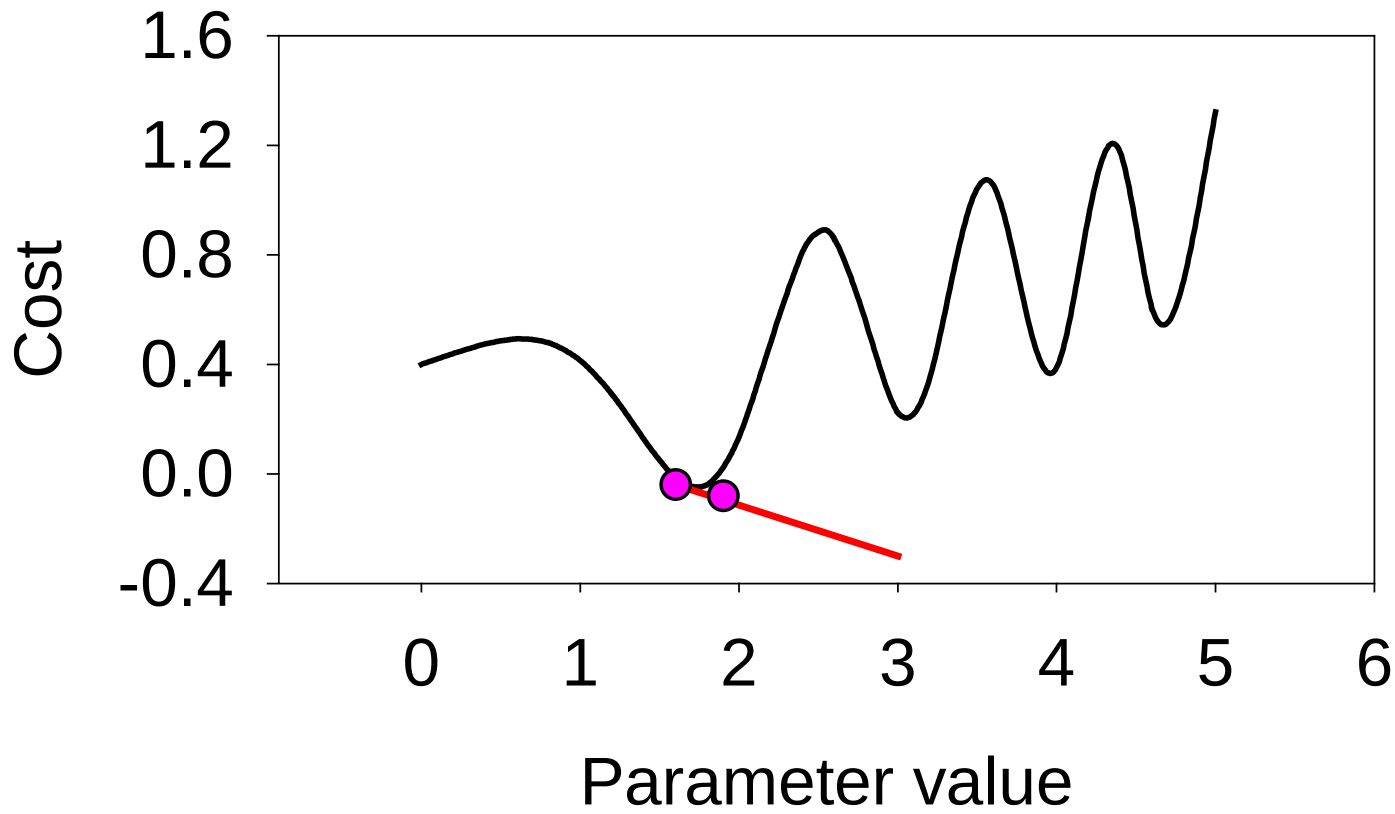
$$\mathrm{cov}(\boldsymbol{b}_{\boldsymbol{es}}) = \sigma^2 \left(\boldsymbol{X}_{\boldsymbol{0}}^{\boldsymbol{T}} \boldsymbol{Q} \boldsymbol{X}_{\boldsymbol{0}} \right)^{-1}$$

- Nonlinear:

$$\mathrm{cov}(\boldsymbol{p}_{\boldsymbol{es}}) = \sigma^2 (\boldsymbol{J}^{\boldsymbol{T}} \boldsymbol{Q} \boldsymbol{J})^{-1}$$

- Solution to nonlinear problem is an tangent linear approximation

Nonlinear parameter estimation



Case study: Penman-Monteith equation

- The equation:

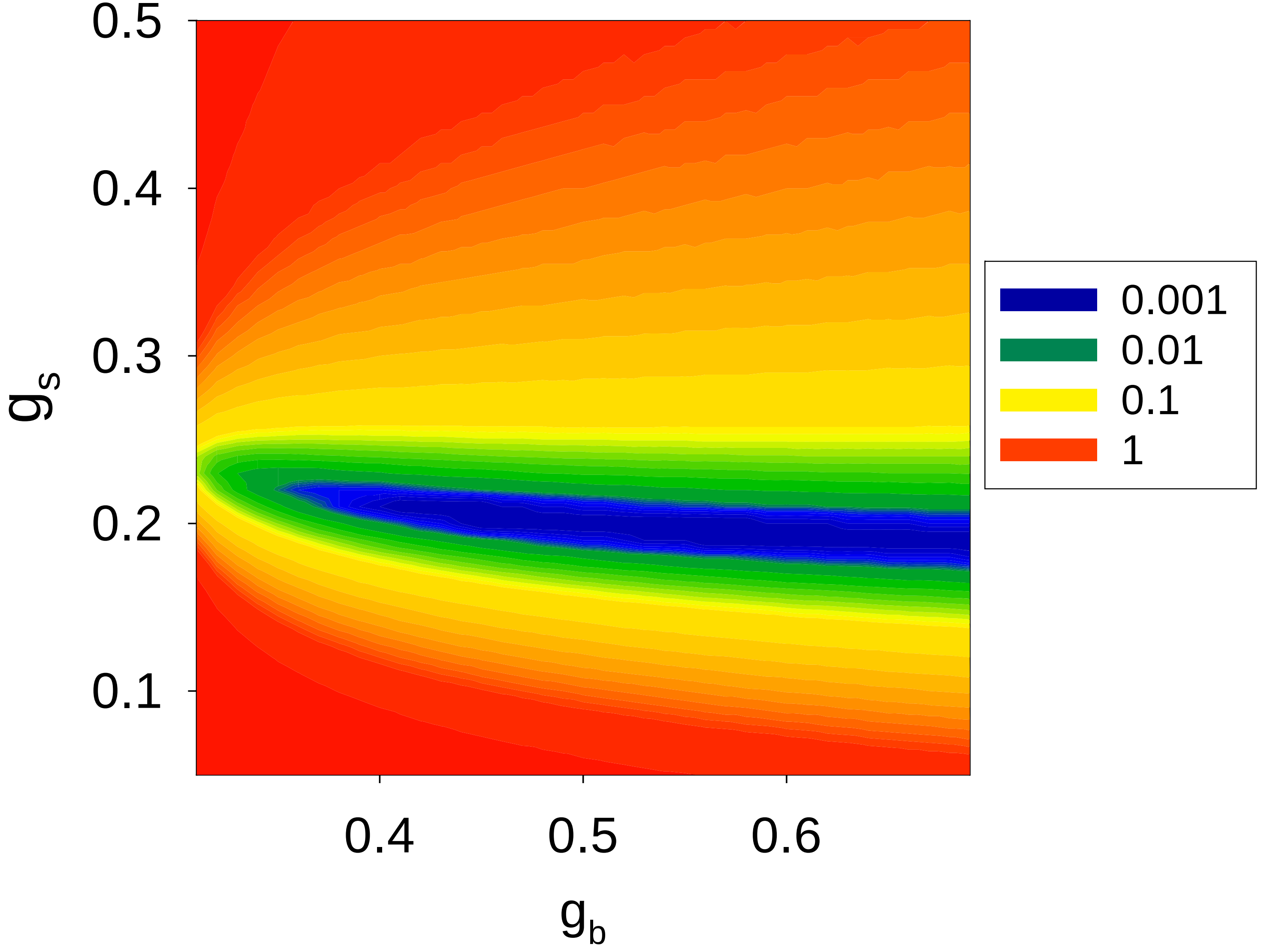
$$\lambda E_f = \frac{sR_{n,i} + D_a c_p \rho_a g_h}{s + \gamma \dfrac{g_h}{g_w}}$$

- Independent variables: T_a, D_a, R_{ni}
- Dependent variable: E_f
- Parameters, g_a, g_b, g_s

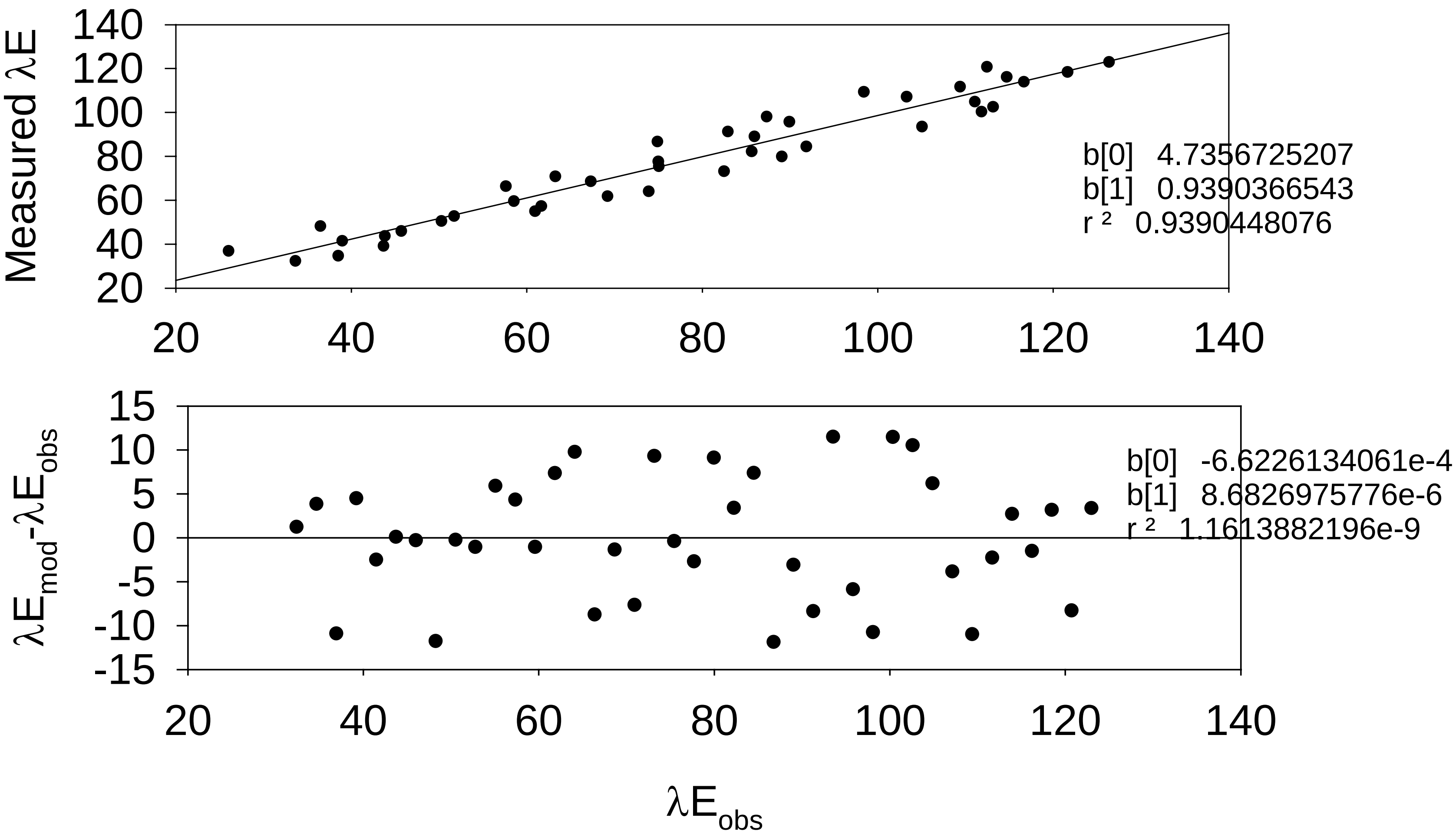
$$g_h^{-1} = g_a^{-1} + \left(0.5 / g_b\right)$$

$$g_w^{-1} = g_s^{-1} + \left(1.075 g_b\right)^{-1} + g_a^{-1}$$

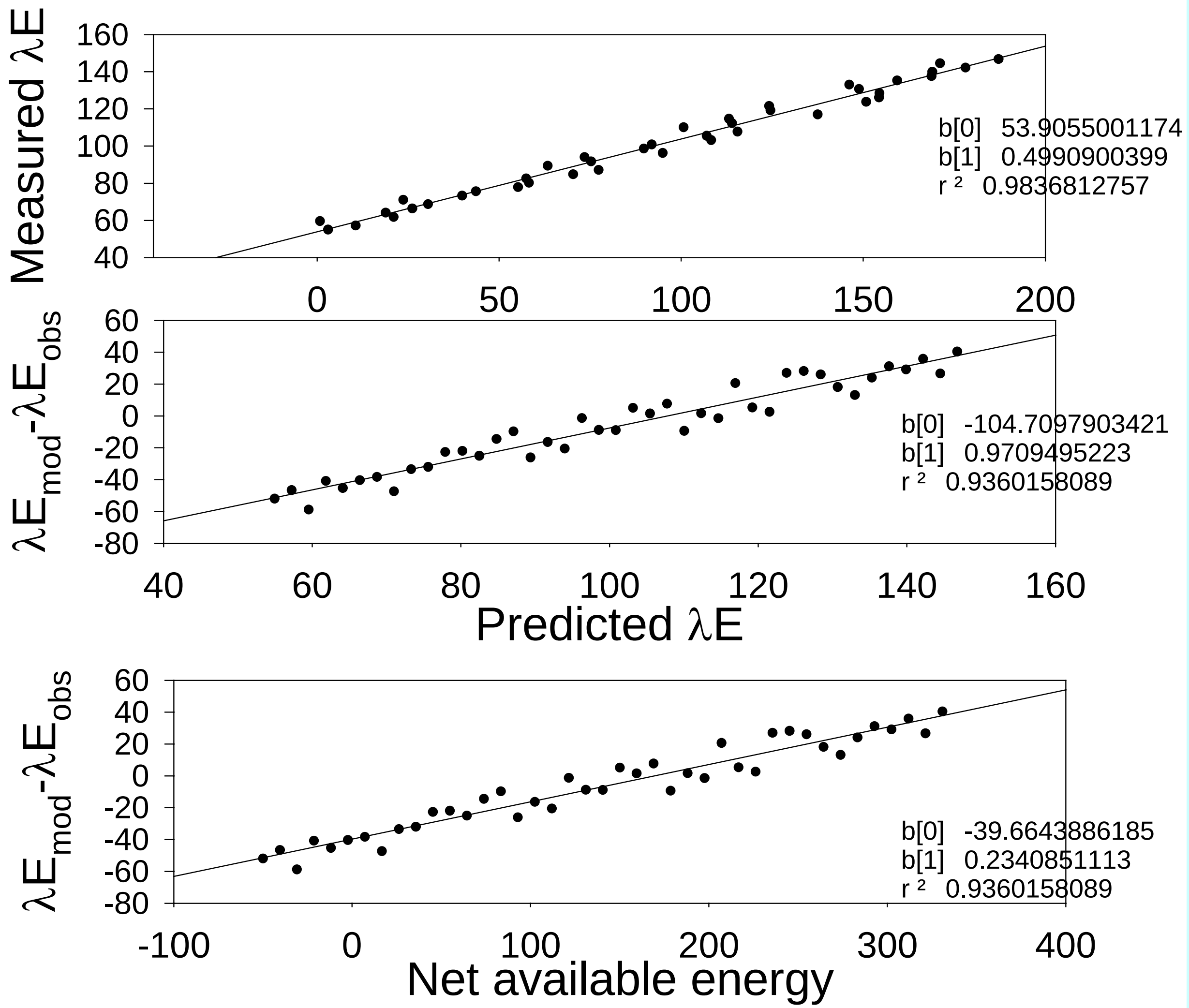
$$\text{Cost} = \frac{1}{n}\sum\left(\lambda E_{\text{mod}} - \lambda E_{obs}\right)^2$$



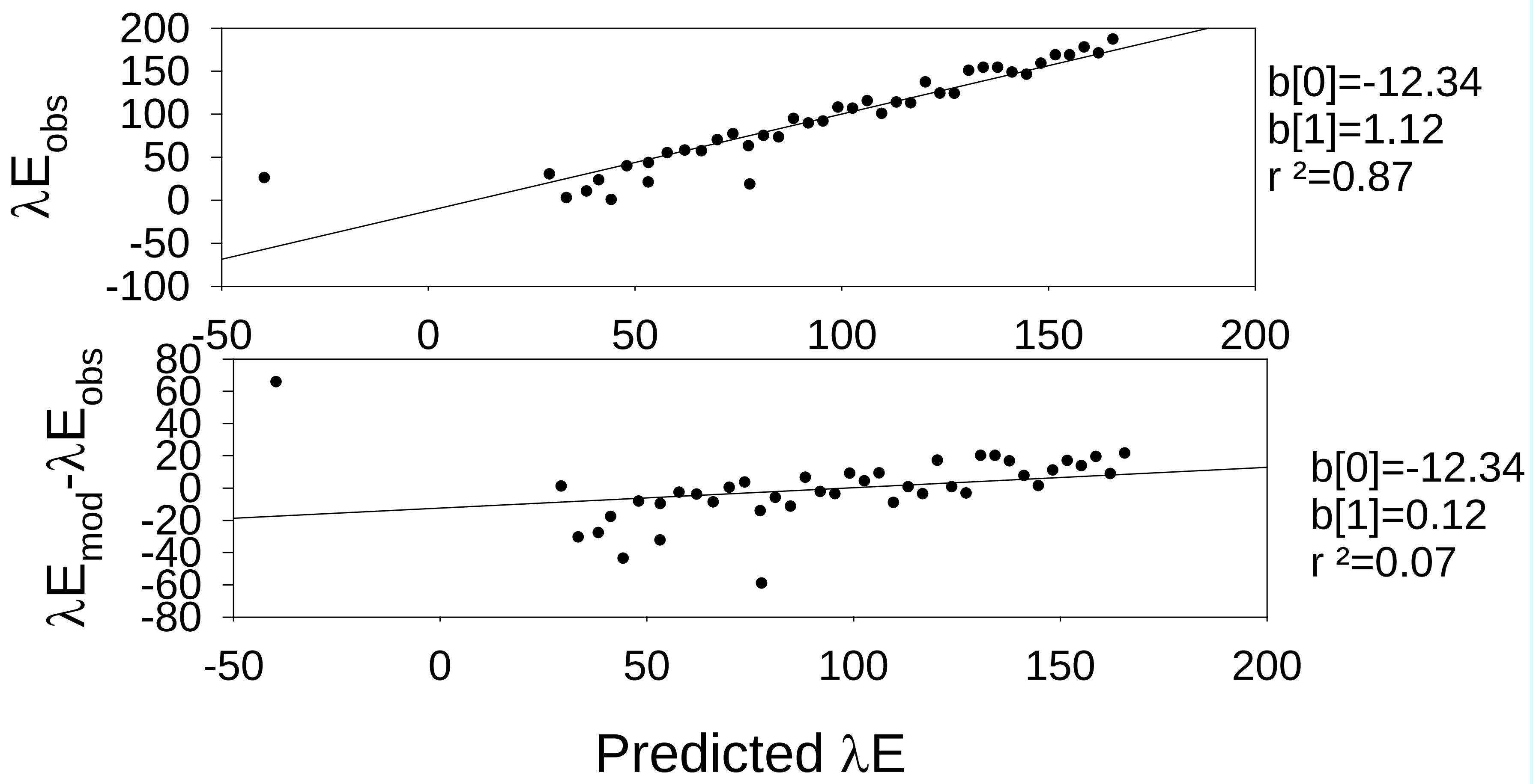
Examining the results



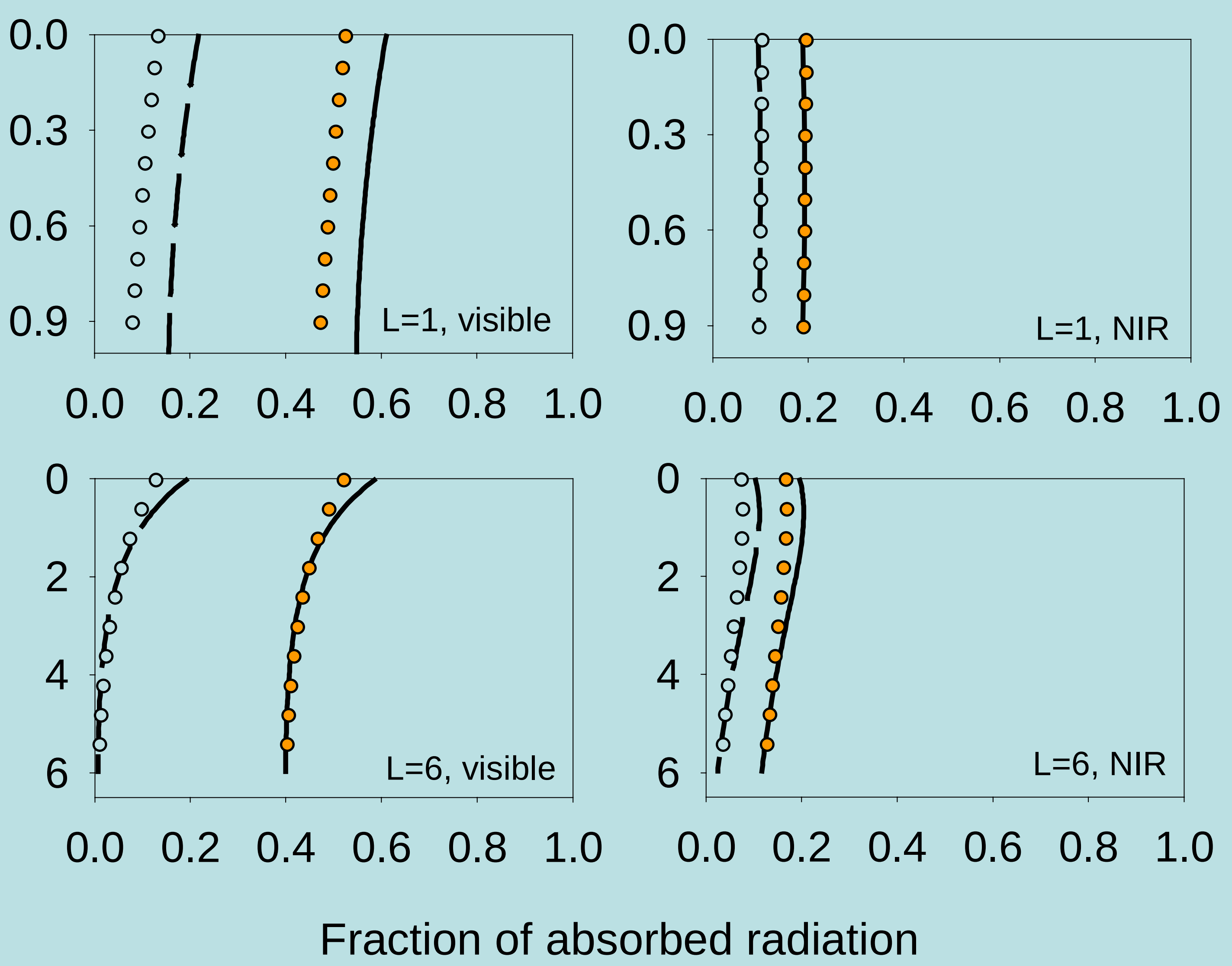
Examining results (case 41)



Examining the results (case 4)



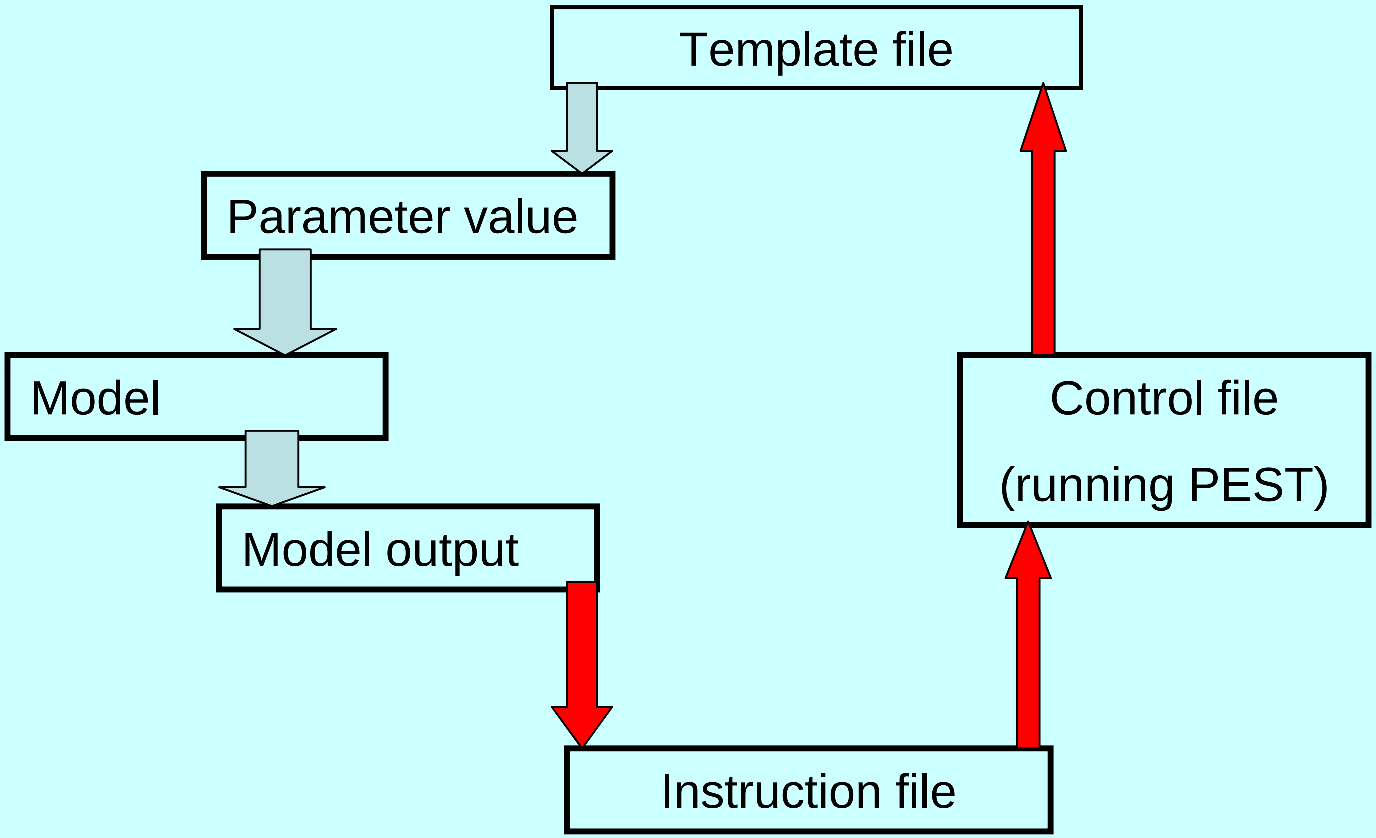
Radiation absorbed in a canopy



Introduction to PEST

- PEST is model-independent, nonlinear parameter estimation package. It is a widely used, free download software.
- It provides stable solution to most nonlinear inversion problems, with the capability of powerful predictive analysis and regularization.
- It communicates to users and models by text files that can be modified by users

How PEST works?



```
pcf
* control data
RSTFLE PESTMODE
NPAR NOBS NPARGP NPRIOR NOBSGP
NTPLFLE NINSFLE PRECIS DPOINT NUMCOM JACFILE MESSFILE
RLAMBDA1 RLAMFAC PHIRATSUF PHIREDLAM NUMLAM
RELPARMAX FACPARMAX FACORIG
PHIREDSWH
NOPTMAX PHIREDSTP NPHISTP NPHINORED RELPARSTP NRELPAR
ICOV ICOR IEIG
* parameter groups
PARGPNME INCTYP DERINC DERINCLB FORCEN DERINCMUL DERMTHD
(one such line for each of the NPARGP parameter groups)
* parameter data
PARNME PARTRANS PARCHGLIM PARVAL1 PARLBND PARUBND PARGP SCALE OFFSET DERCOM
(one such line for each of the NPAR parameters)
PARNME PARTIED
(one such line for each tied parameter)
* observation groups
OBGNME
(one such line for each observation group)
* observation data
OBSNME OBSVAL WEIGHT OBGNME
(one such line for each of the NOBS observations)
* model command line
write the command which PEST must use to run the model
* model input/output
TEMPFLE INFLE
(one such line for each model input file containing parameters)
INSFLE OUTFLE
(one such line for each model output file containing observations)
* prior information
PILBL PIFAC * PARNME + PIFAC * log(PARNME) ... = PIVAL WEIGHT OBGNME
(one such line for each of the NPRIOR articles of prior information)
```

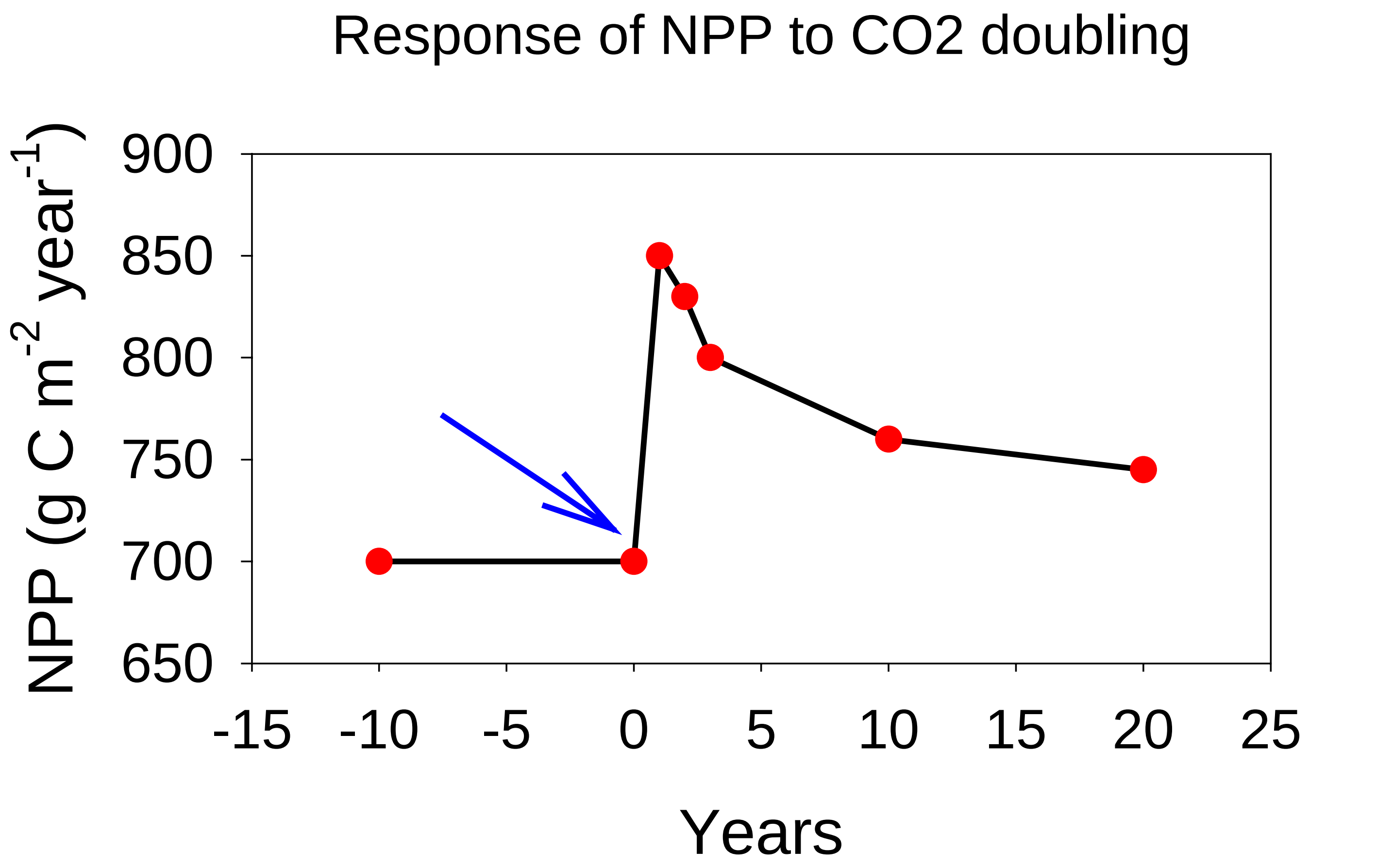
Template file

```
ptf #
      ratioRL        resprc1           tempcoef1               tempcoef2
#     ratioRL  #,#   resprc1   #,#     tempcoef1       #,#     tempcoef2      #
```

Instruction file

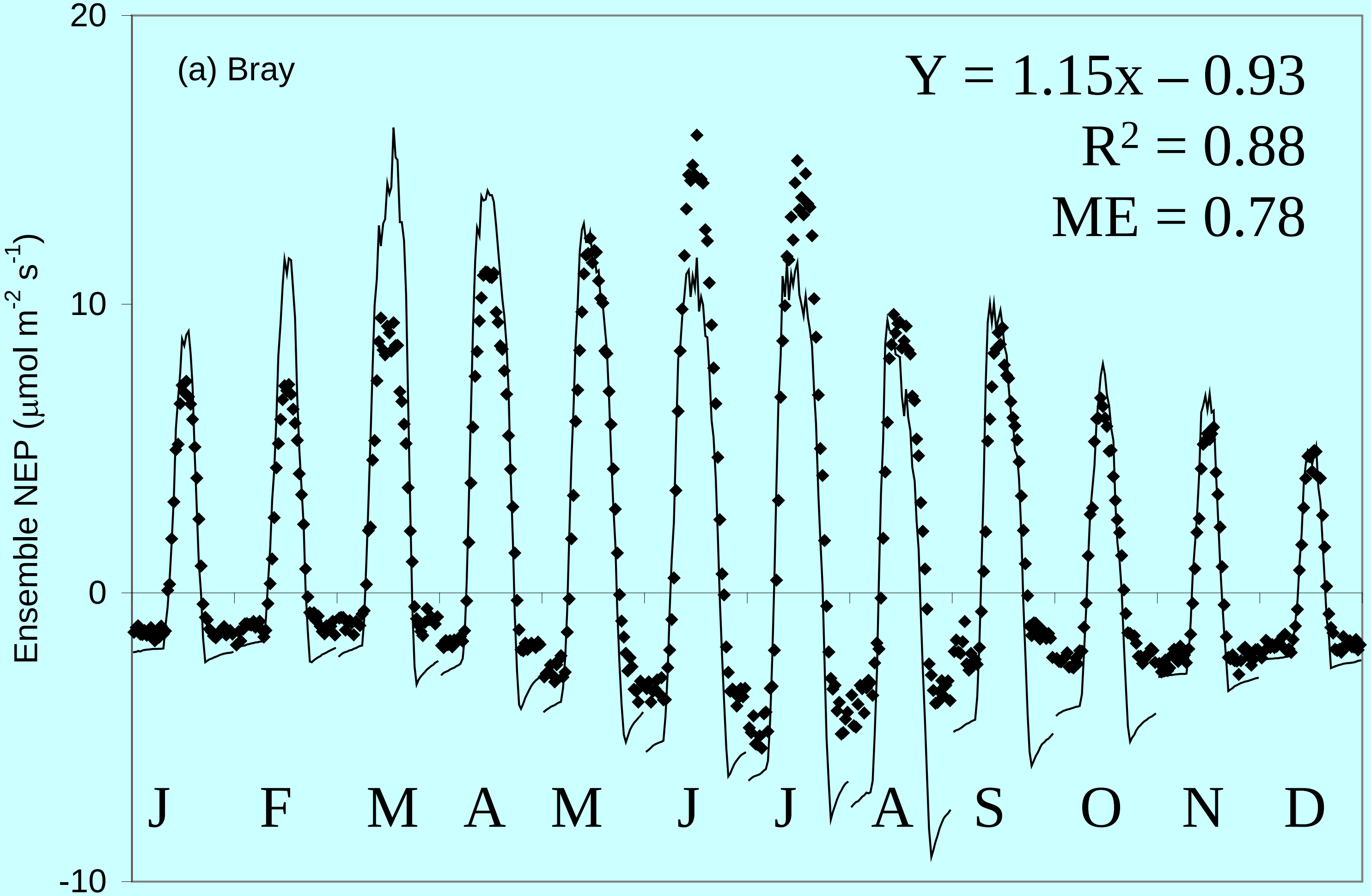
```
pif #
L1  (A0000001)13:22
L1  (A0000002)13:22
L1  (A0000003)13:22
L1  (A0000004)13:22
L1  (A0000005)13:22
L1  (A0000006)13:22
L1  (A0000007)13:22
L1  (A0000008)13:22
L1  (A0000009)13:22
L1  (A0000010)13:22
L1  (A0000011)13:22
L1  (A0000012)13:22
L1  (A0000013)13:22
L1  (A0000014)13:22
L1  (A0000015)13:22
L1  (A0000016)13:22
L1  (A0000017)13:22
L1  (A0000018)13:22
L1  (A0000019)13:22
L1  (A0000020)13:22
L1  (A0000021)13:22
L1  (A0000022)13:22
```

Application I: interpretation



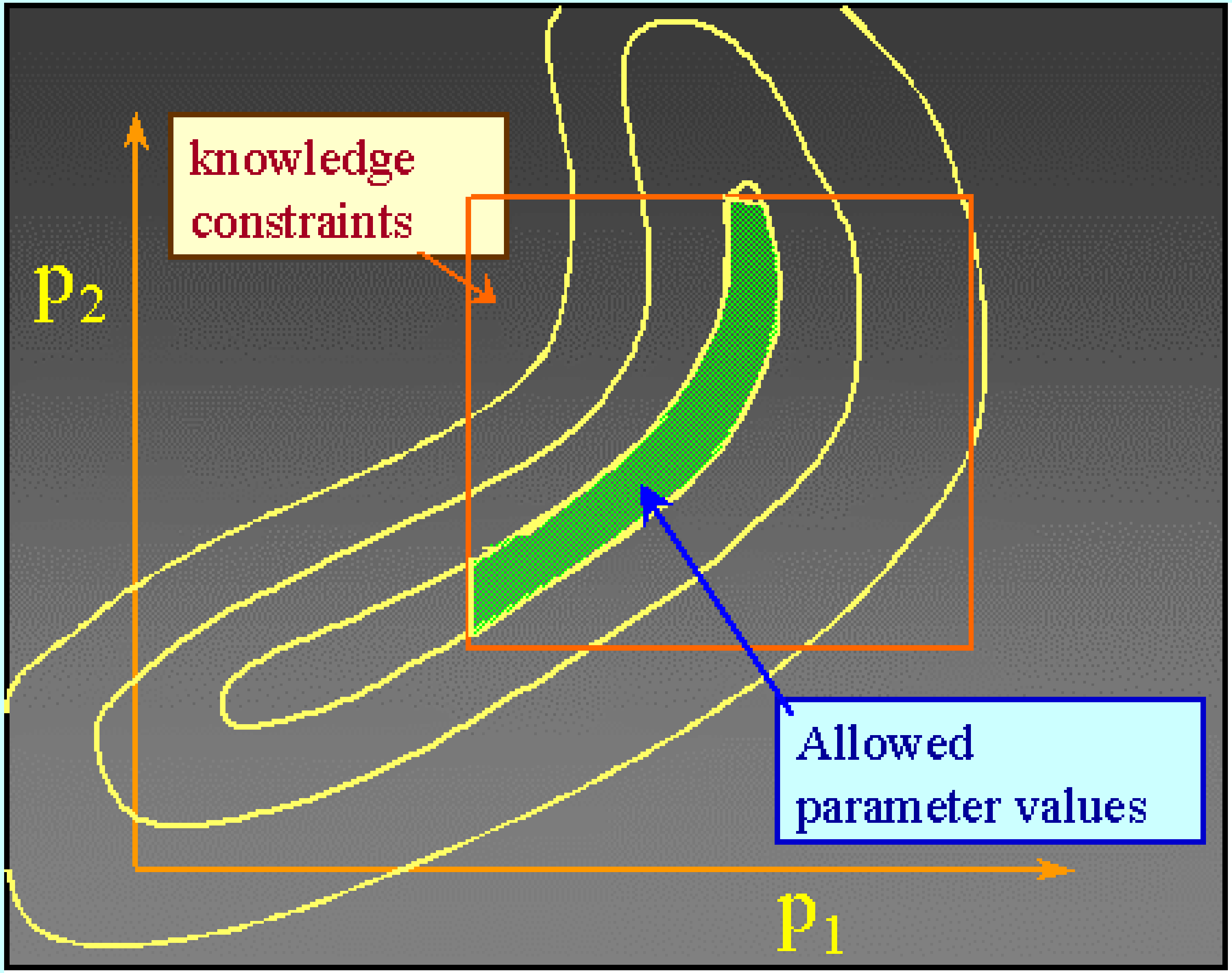
See Wang, McMurtrie, Medlyn and Pepper (2006)

Application II: calibration

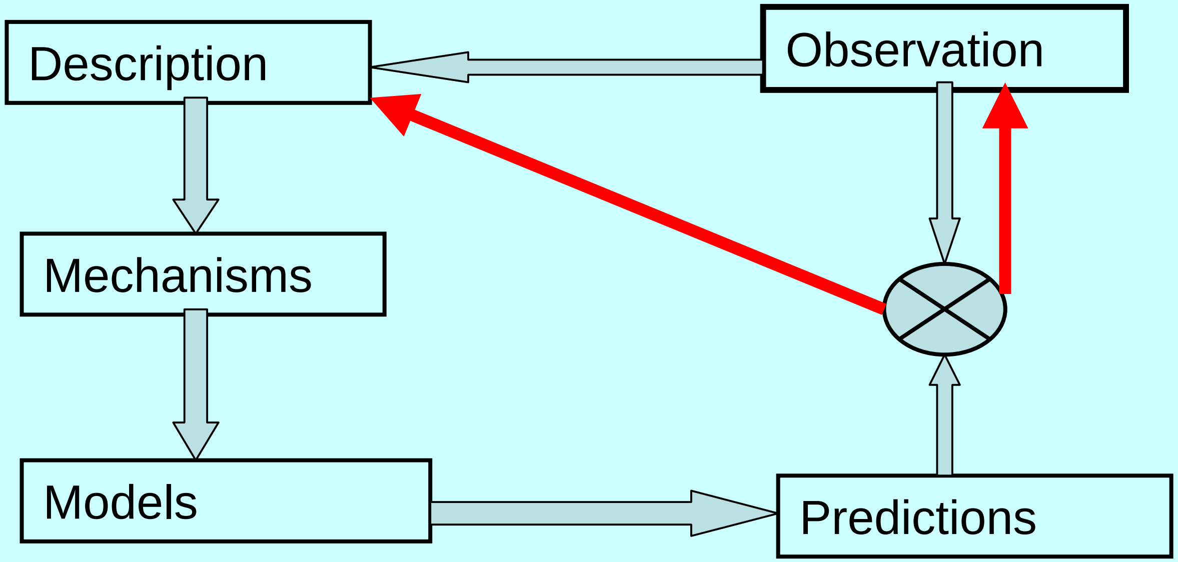


Medlyn et al. 2005

Application III: predictive analysis



Model as a set of hypotheses



Land surface modeling

- As a key component of earth system
- Synthesis leads to a better representation of ecological processes
- Use data to test the model
  - Discrepancies lead to improvement, new discoveries

What are we doing differently from 10 years ago?

- More measurements with greater temporal and spatial resolution
- More synthesis and generalization
- Uncertainties in model and observations

Model-data fusion

- A technique being applied in physical sciences since 1960's, and will become a standard tool within the next decade or so;
- A platform to facilitate the interactions between modelling and observations
- Synthesis of information of different scales

Future

- Model data fusion as a power data synthesis tool

- Emergence of global change biology and earth system sciences

- Reducing uncertainties