CAS-CEOP-CAHMDA
2010, Lahsa

Approaching the 'Ground Truth'

Jiemin Wang

Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences
jmwang@lzb.ac.cn

Data-Model Fusion

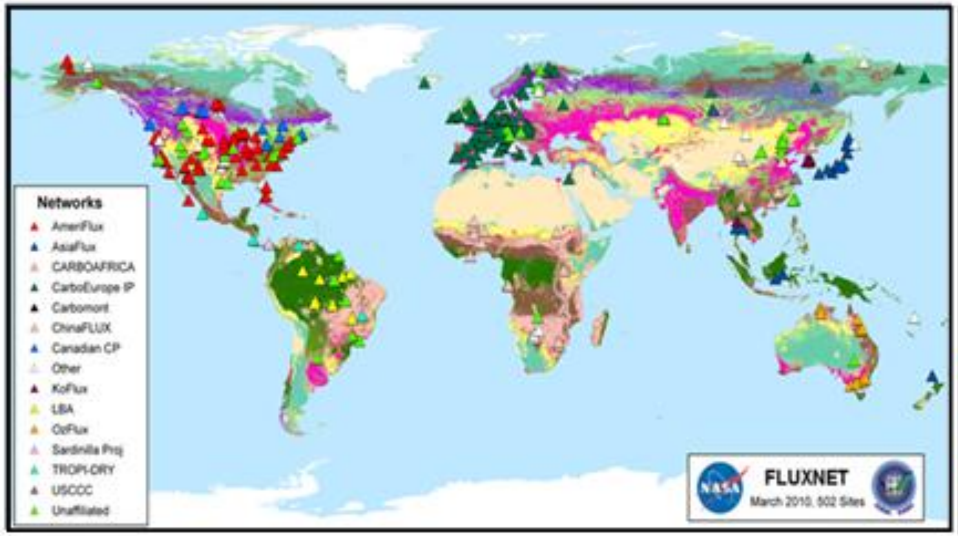
Challenges:

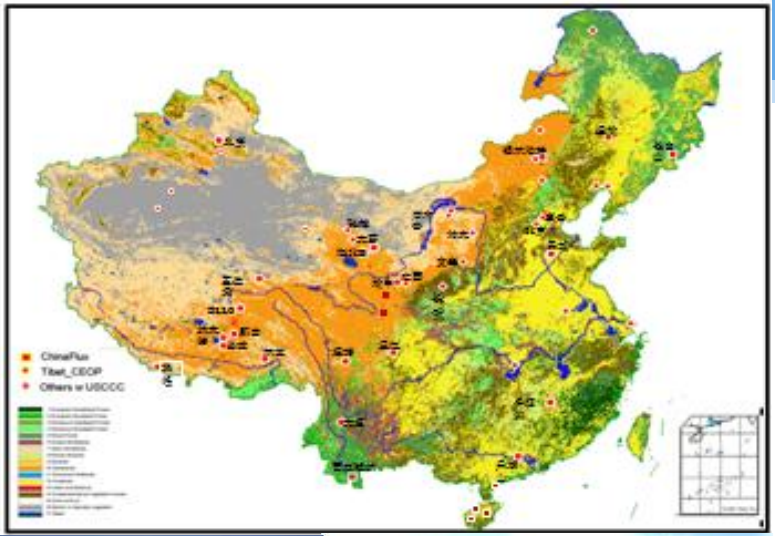
- ***'Data are not truth'***
  (main subject of this presentation)
- ***'No model is perfect'***
  - Model is based on theories. Any theory is conditional.
  - Model needs proper inputs and parameterization schemes. All has uncertainties.
  - Model needs validation, which is not an easy procedure.
- ***Good data + Good modeling***
  ⇨ ***Good Prognostication***

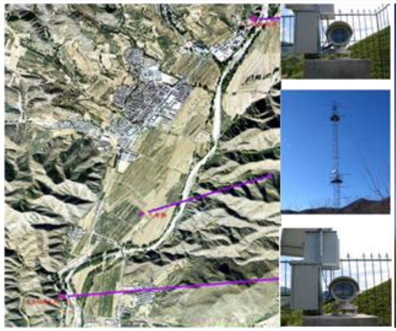
Challenges for ‘*Good Data*’

- **Accuracy & Representativeness** for *Primary variables*: Wind, Temperature, Humidity, Pressure, Radiation, precipitation,…
  - Errors of observation, systematic and/or random
  - Representativeness, spatial & temperal
- **Processing schemes** for *Derivatives*: Roughness length, Emissivity, Surface temperature, LAI, Resistances, Structure parameter, u_*, Heat flux, Evapotranspiration, CO_2 flux, …
  - Complexity of these schemes vary with different applications
- **Inconsistencies between multiple data streams.** One parameter may be observed by various sensors, or, derived with different methods.

Have we got what we want from these stations?







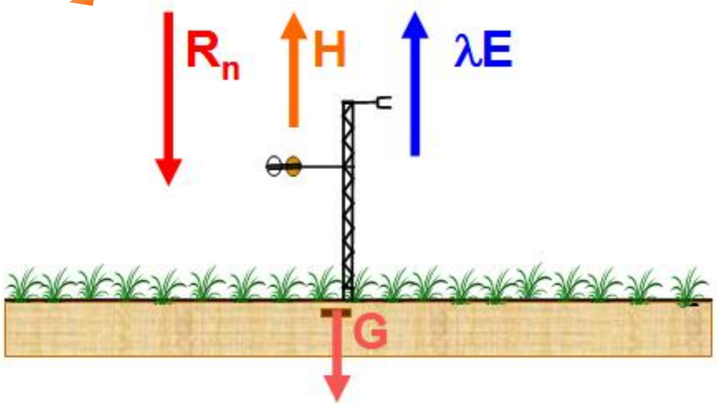
Advanced instruments

- **Advanced instruments do not guarantee the data accuracy**
- **Advanced instruments are normally with complex theories behind.**
  - Any theory is based on some conditions.
- **Specific knowledge is needed in**
  - Site selection, sensor mounting, station operation, maintenance, etc.
  - Data processing
  - Quality Control/ Quality Assessment.

Observations in flux stations

- Basic fluxes measurements :
  - **Tower** (profiles of wind, temperature, humidity, …);
  - **Eddy-covariance fluxes (momentum, heat, water vapor, CO_2, etc.)**
  - **Radiation** components;
  - **Soil** (temperature, water content, heat flux, etc.);
  - **Large aperture scintillometer** (**LAS**)
- Specific *hydrological*, *ecological*, *aerosol*, and *surface remote sensing* observations

The surface energy balance closure problem



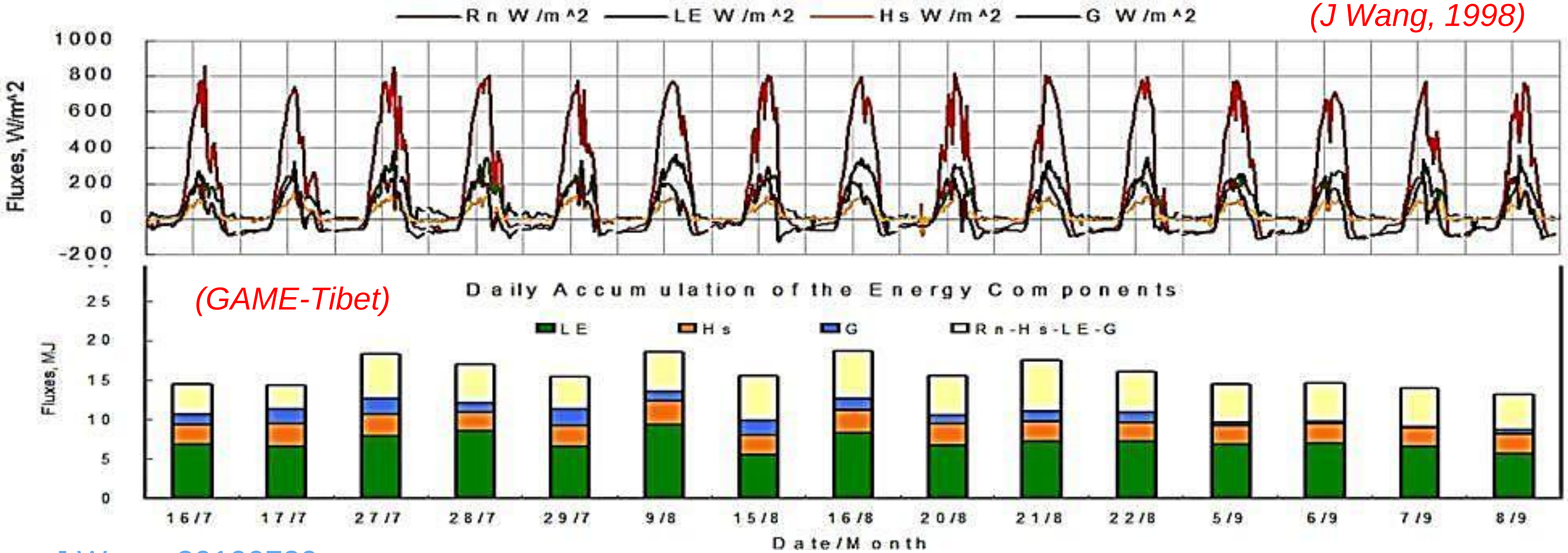
$$R_n - G = H + \lambda E \quad ?$$

$$Res = R_n - G - H - \lambda E \approx 50 \sim 300 \ (W/m^2)$$

$$CR = (H + \lambda E)/(R_n - G) \approx 60\% \sim 90\%$$

H & λE are from Eddy-covariance method

(J Wang, 1998)



Eddy-covariance flux system

- Fluctuations measured: $u, v, w, T_S, q, C, \ldots \; (C = \overline{C} + C')$
- Fluxes: $\tau = \rho \overline{u'w'}, \quad H \approx \rho C_p \overline{w'T_S'}, \quad \lambda E \approx \lambda \overline{w'q'}, \quad F_C \approx \overline{w'C'}, \ldots$
- A simplification from rather complex equation for scalar C

$$\int_0^{z_r} S(t,z)\,dz = \int_0^{z_r} \frac{\partial \bar{c}(z)}{\partial t}\,dz + \overline{w'c'}(z_r) + \int_0^{z_r} \bar{w}(z)\frac{\partial \bar{c}(z)}{\partial z}\,dz + \int_0^{z_r} \left(\bar{u}(z)\frac{\partial \bar{c}(z)}{\partial x} + \bar{v}(z)\frac{\partial \bar{c}(z)}{\partial y} \right) dz$$

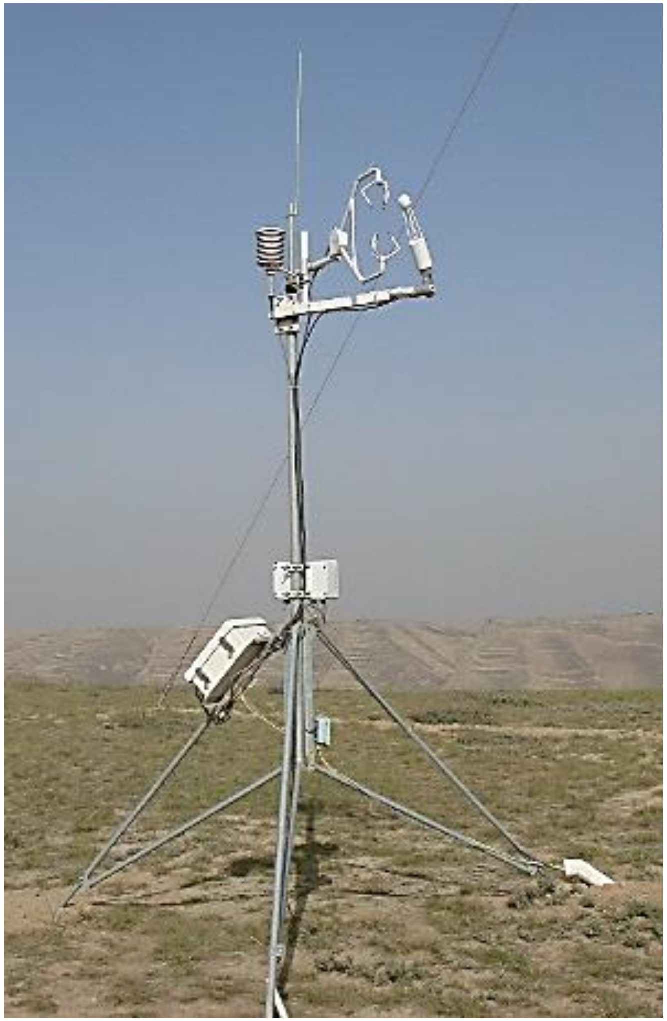
Source / Sink | Storage | Turbulent flux | Vertical advection | Horizontal advection

$$\Rightarrow \int_0^{z_r} S(t,z)\,dz \approx \overline{w'c'}(z_r)$$

Source / Sink | Turbulent flux

With assumptions:
- Fully developed turbulence
- Stationary
- No advection, …

Careful post-field data processing is indispensable



u, v, w, T_s, q, C,….

↓

Remove anomalies

↓

Tilt Correction (Pf / 2D rotation)

↓

Frequency response correction

↓

Density (WPL) correction

↓

Sonic virtual temp. correction

↓

Fluxes: τ, H, λE, F_C,…

↓

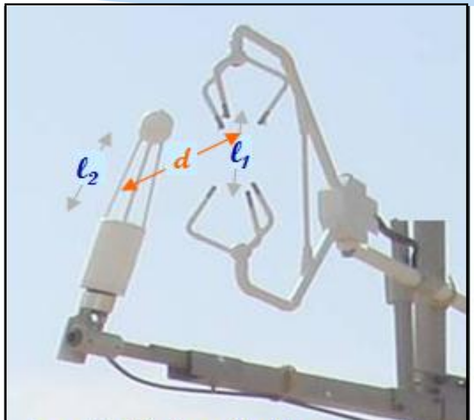
Quality assessment

High frequency lose due to sensors path & separation

$$\overline{w'x'} = \int_0^\infty Co_{wx}(f)df$$

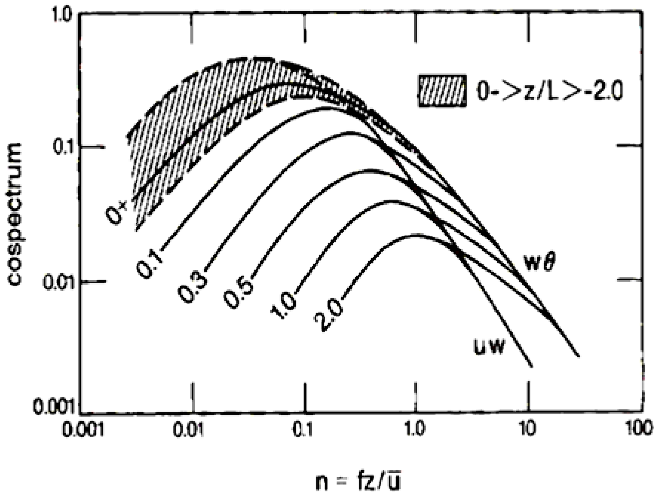
$$\left(\overline{w'x'}\right)_m = \int_0^\infty T(f)\cdot Co_{wx}(f)df$$

$$T(f) = T_{l_1}(f)\cdot T_{l_1}(f)\cdot T_d(f)\cdot\ldots$$



Length ℓ & d resulted 'line-average'. Sensible eddy >2d.

Standard co-spectra of uw & wθ



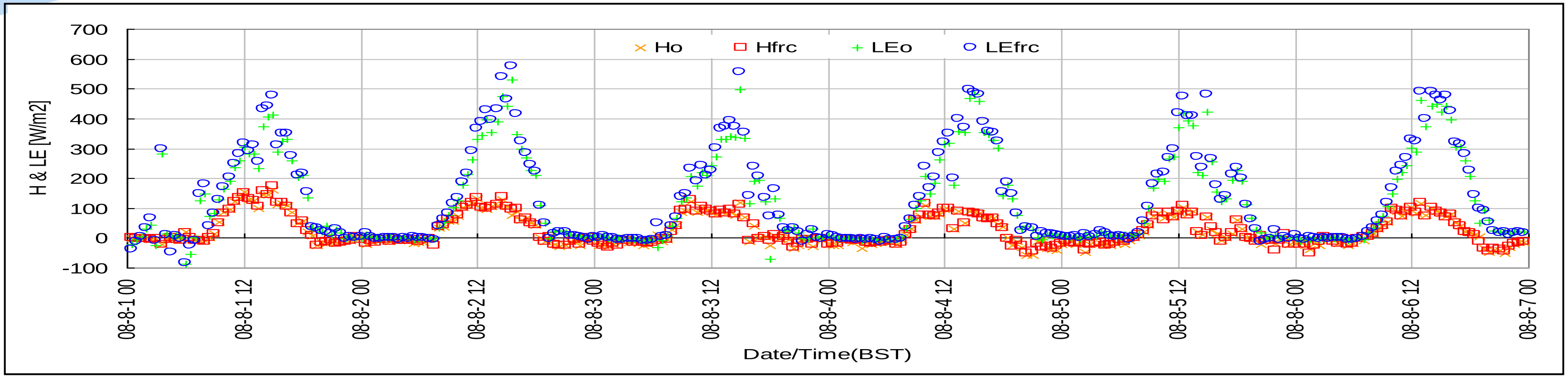
'Standard' co-spectra

$$-\frac{fCo_{uw}(f)}{u_*^{\,2}} = 0.05n^{-4/3}G(z/L)$$

$$-\frac{fCo_{wT}(f)}{u_*T_*} = 0.14n^{-4/3}H(z/L)$$

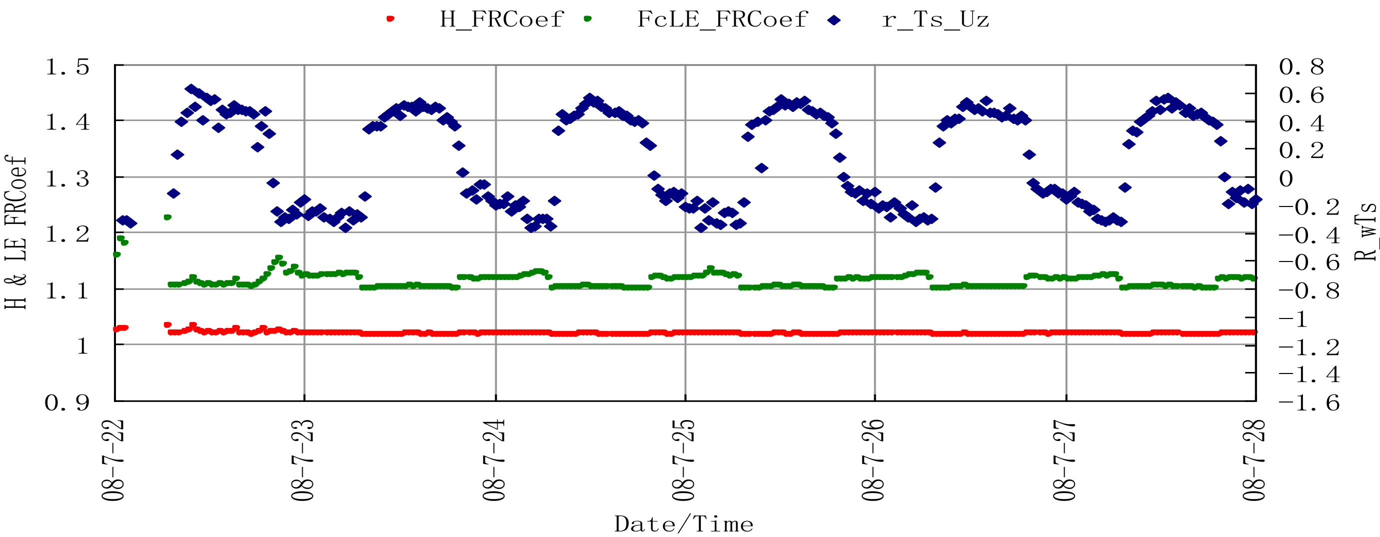
Some spectral models can be used to correct limited average period induced low-frequency lose.

Comparison of frequency corrected fluxes (Arou, Aug 1-6, 2008)



Correction coefficients for H & LE (& Fc)

Note: the change of correlation r_{wTs}, an indicator for low freq. effects.



Surface energy imbalance:- ***Multiple reasons***

I. Specific points for Eddy-Covariance flux system
   - Proper data reprocessing procedure & QA/QC
   - The 'inability' of EC system in sampling slow moving TOS?

II. Accuracy of relevant instruments

III. Different source areas for different sensors

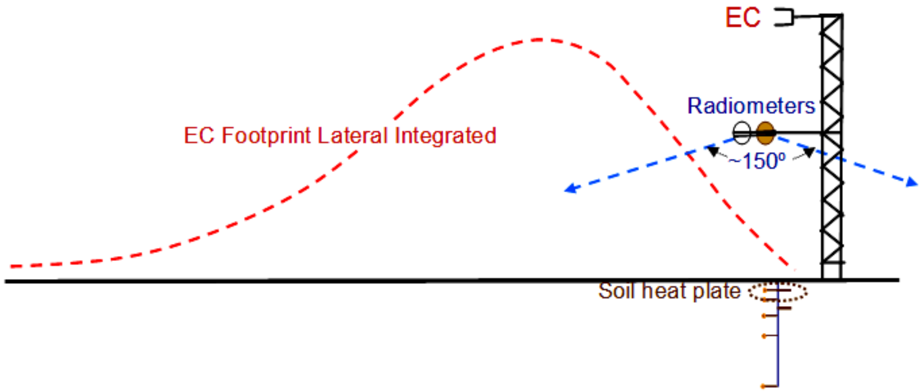
IV. Additional components in surface energy budget
   - Soil heat storage in upper soil layer
   - Heat storage in vegetation canopy and in the air layer
   - Photosynthetic consumption during daytime
   - Humidity change, dew formation, etc.

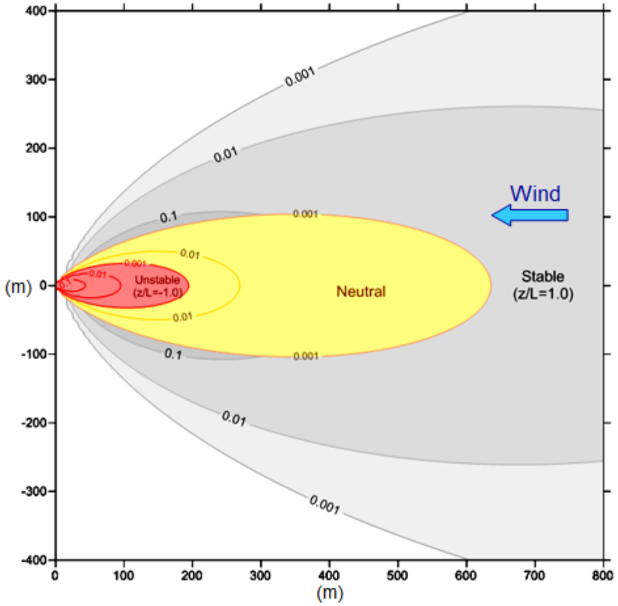
V. Effects of horizontal/vertical *advection*

Source areas for different sensors

	Observation height (m)	Horizontal scale of source area (m)
EC fluxes	2-10	100~1000
Radiometers	1.5-2	10~15
Soil heat flux	-0.02 ~ -0.1	0.1

(Foken, 2006)





Source area for flux observation *changes with sensor height (z), wind speed (U), wind direction, atmospheric stability (z/L), etc.*

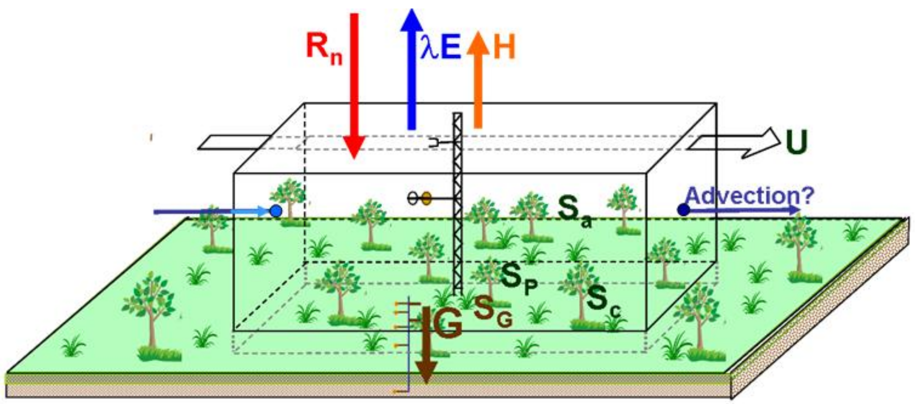
left: *Footprint calculation for case:-*

z = 3.0 m, U =1.5 m/s, z_0 *= 0.03, stdv =* 0.3, *z/L=*[*-1.0, 0, 1.0*]

by Kormann & Meixner (2001) method

- Red: unstable (*z/L*= -1.0); Yellow: neutral; Grey: stable (*z/L*=1.0)
- Contour lines: outer, 0.001; middle, 0.01, inner, 0.1 (of the max footprint value)
- For the flux contribution reduces to the 1% of the maximum point, the upper-wind distances for unstable, neutral, & stable cases are about 100 m, 300 m, & 1500 m, respectively.

Additional components in SEB



Additional terms:

- S_p: Photosynthetic flux *(Jacobs et al. 2003; Meyers and Hollinger 2004)*
- S_a: Air enthalpy change *(Atzema, 1993)*
- S_c: Vegetation canopy enthalpy change *(Oncley et al. 2007)*
- Others: S_q (air moisture change), S_d (dew formation), … *(Jacobs et al. 2008)*

- The evaluation of ***surface soil heat flux*** has not been paid enough attention until the 1990's. Methods for evaluate soil heat storage S_G are based on soil temperature (& moisture) & heat flux observation at some depth, e.g. Carloric method *(Fritschen & Gay 1979)*

The contribution of the storage terms to the surface energy balance closure *(Jacobs et al. 2008)*

	S_G	S_p	S_a	S_c	S_d	S_q
Additional improvement	9.0%	3.0%	2.0%	0.5%	0.1%	0.5%

Correction of heat storage in upper soil layer: *Two methods*

- **Harmonic analysis (HM)** (*Heusinkveld et al., 2004*)

Using time series of multi-level observations of soil temperature and at least one level heat flux, then, heat flux at level z:

$$G(z,t)=\lambda\cdot\sum_{k=1}^{M}\left\{A_k C_s\sqrt{k\omega\alpha}\exp\left(-z\sqrt{k\omega/2\alpha}\right)\sin\left[k\omega t+\Phi_k+\left(\pi/4\right)-z\sqrt{k\omega/2\alpha}\right]\right\}$$

- **Temperature prediction & correction (TDEC)** (*Yang & Wang, 2008*)

Using heat diffusion equa., adjusting the temperature profile with observations. Soil flux obtained through integrating the tempera-ture profile, with advantages:

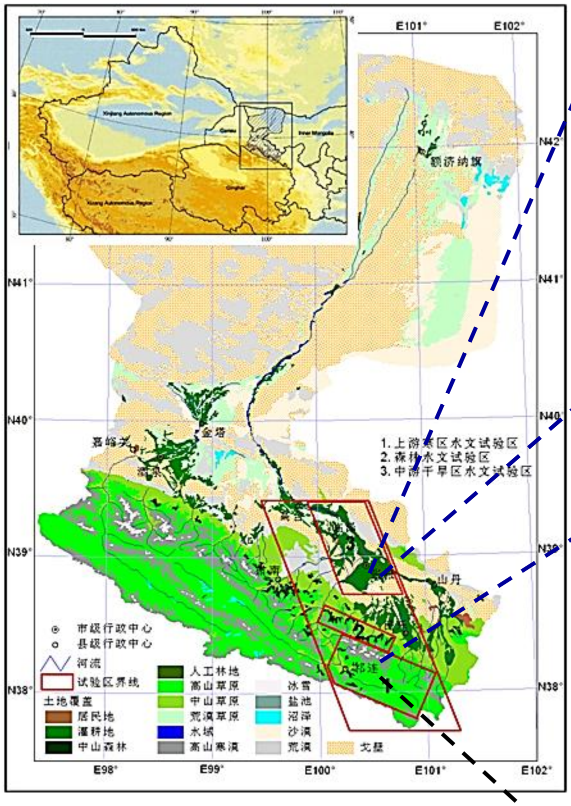
- Assumption of 'soil vertical homogeneity' not needed
- Results not sensitive to the initial value of soil heat conductance λ

$$G(z,t)=G(z_{ref},t)+\frac{1}{\Delta t}\sum_{z_{ref}}^{z}\left[c_v\left(z_i,t+\Delta t\right)T\left(z_i,t+\Delta t\right)-c_v\left(z_i,t\right)T\left(z_i,t\right)\right]\Delta z$$

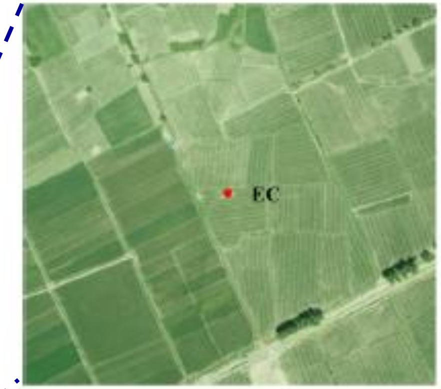
Case study:
Two stations of WATER Project

August 1-6, 2008

WATER Proj.
(Heihe River Basin)

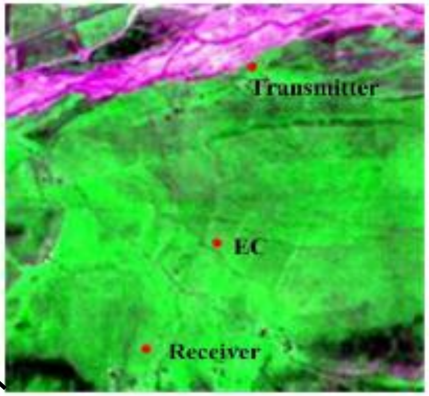


Yingke
(Oasis cropland)



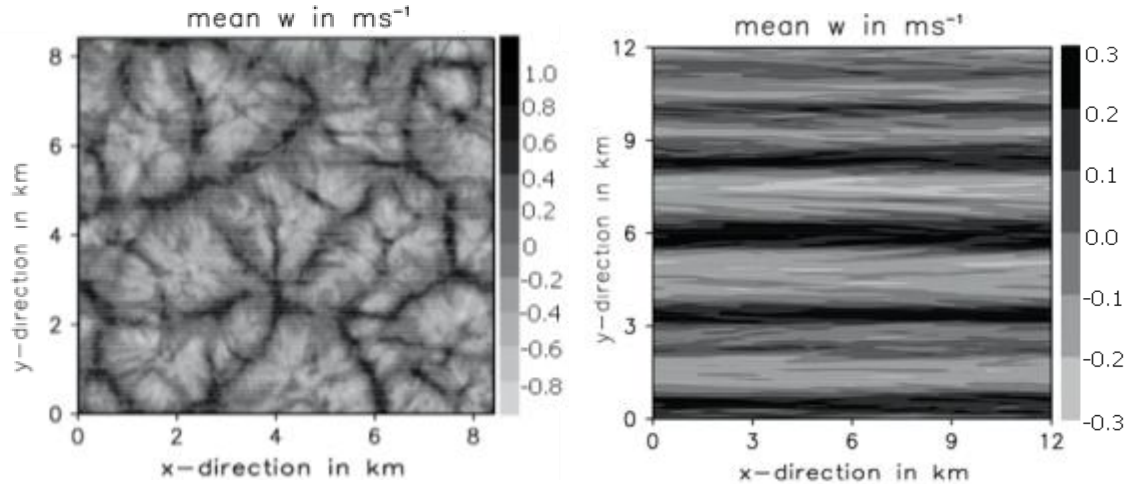
- Former, CR ≈ 71%
- with frequency response corr.
  CR ⇒ 82%
- with soil storage corr.
  CR ⇒ 89%
- with canopy photosynthesis corr.
  CR ⇒ 92%.
- with Air/canopy h. storage corr.
  CR ⇒ 93%

Arou
(Short grass prairie)



- Former, CR=85%
- with frequency response corr.
  CR ⇒ 92%
- with soil storage corr. (TDEC)
  CR ⇒ 98%

Turbulence Organized Structures (TOS) from Large Eddy Simulation (LES)

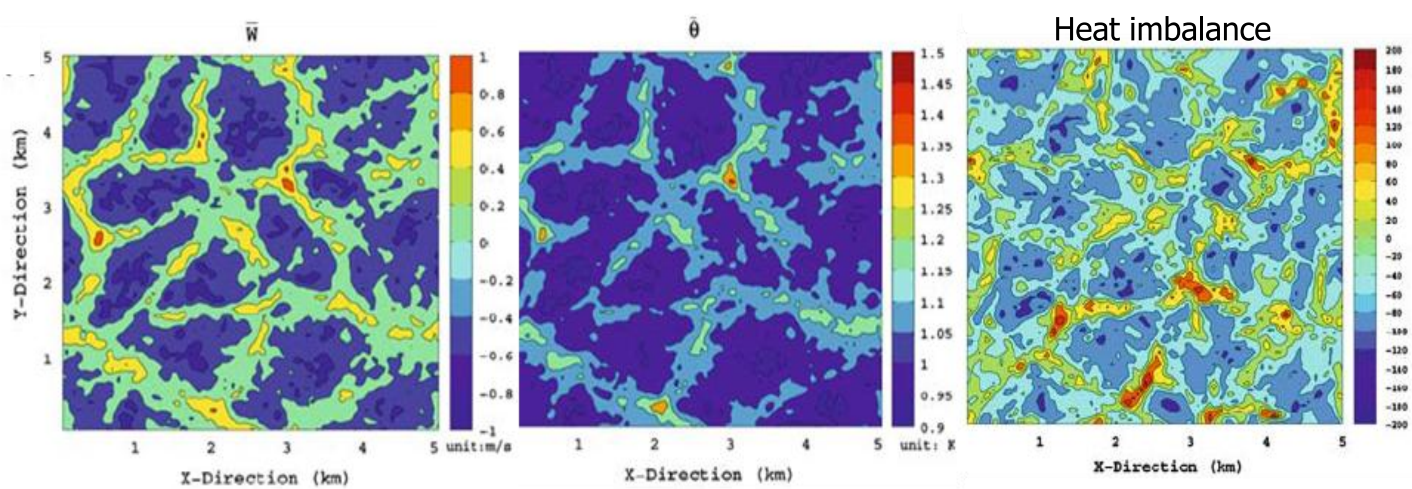


Vert. wind maps
(z=100m, 1hr mean)

Left： Ug= 0 m/s
Cell-type convection

Right: Ug=15m/s
Roll-like convection

(Steinfield et al, 2007)



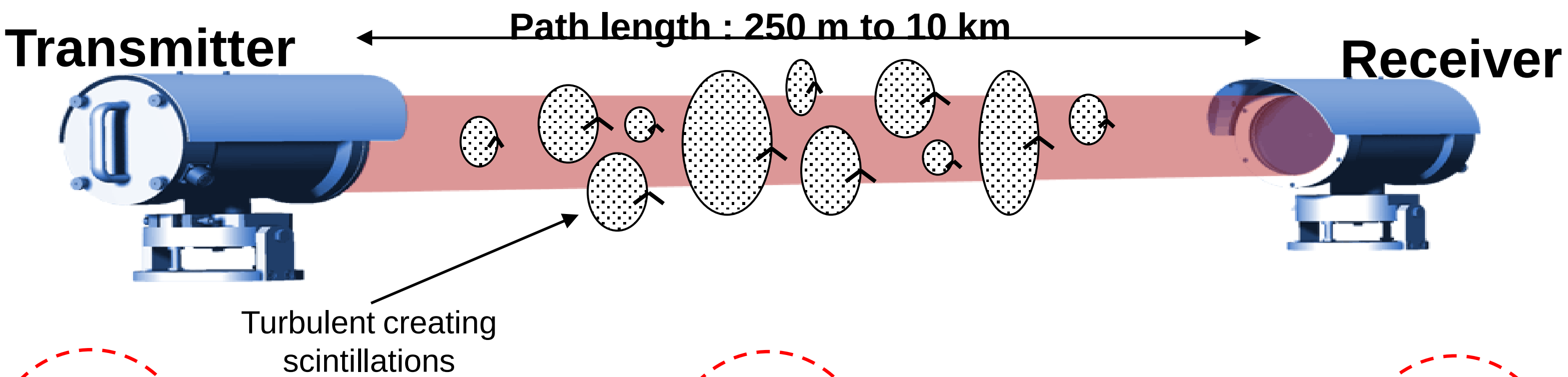
w, θ, & imbalance.

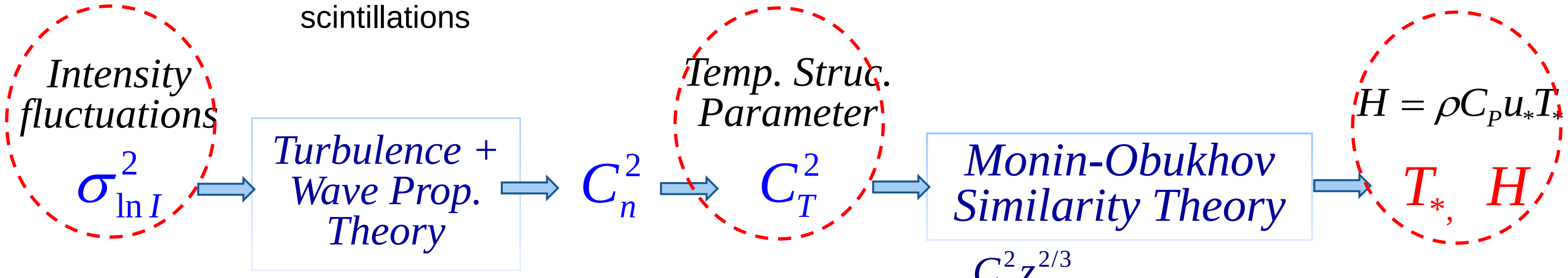
Simulated at Ug= 0 m/s, z=100m, 1.8~2.8hr mean

(Huang et al, 2008)

Large Aperture Scintillometer (LAS)

Comparing with ***EC:*** *Advantages -* ***Larger area coverage***
Disadvantage - ***MOST is needed***



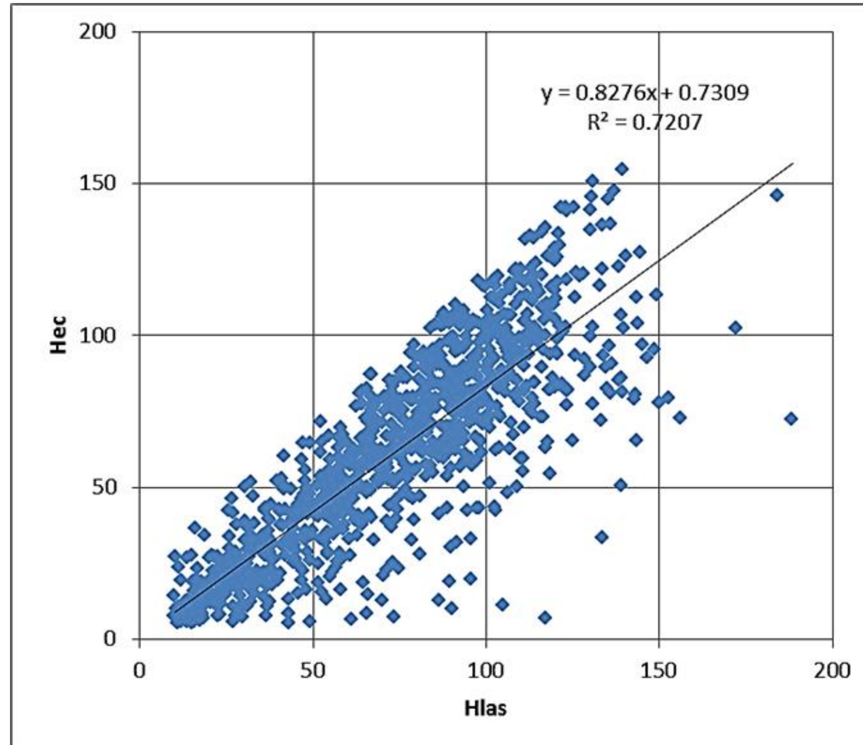


$$\frac{C_T^2 z^{2/3}}{T_*^2} = f_T(z/L)$$

e.g. for z/L<0:- $f_T(z/L) = c_{T1}(1 - c_{T2}\ z/L)^{-2/3}$

Large Aperture Scintillometer

Arou, Heihe Basin. Effective height: EC, 3.15m, LAS, 9.5m



$$\left\langle \left(H_{LAS} - H_{EC} \right) / H_{LAS} \right\rangle = 16.7\%$$

Comparison of sensible heat flux from LAS and from EC (Arou, Jun-Aug, 2008).

The large difference seems unrealistic.

LAS: Similarity Functions

Popular similarity function used:

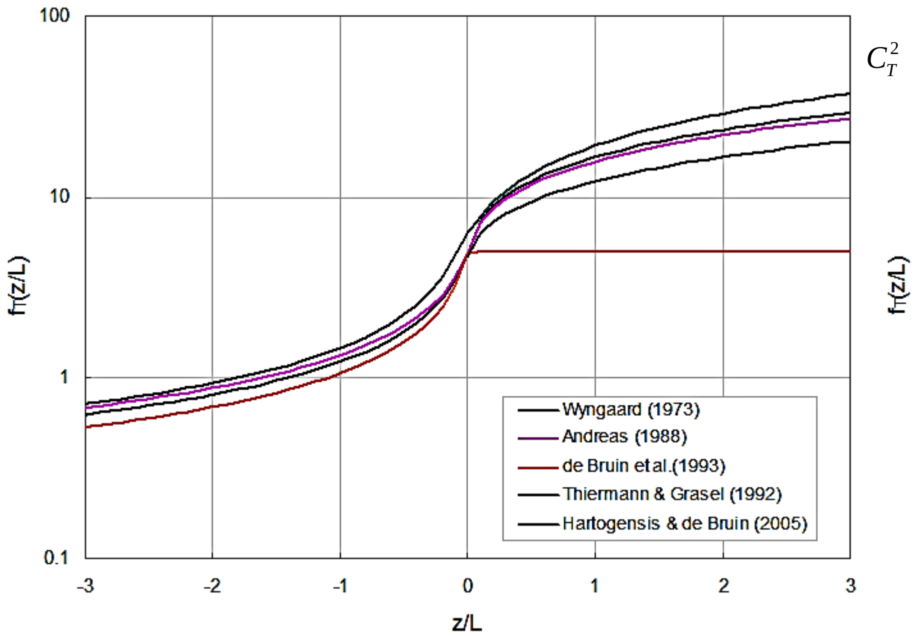
Wyngaard (1973), Andreas (1988), De Bruin et al. (1993), Thiermann & Grassl (1992), Hartogensis & de Bruin (2005)

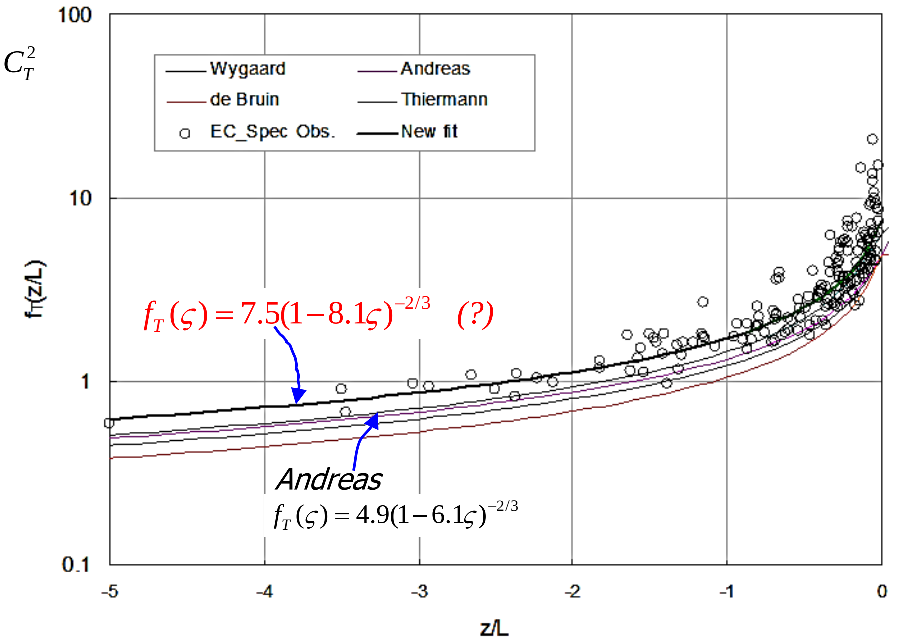
Relative difference is ~ 10-15%, more in stable conditions

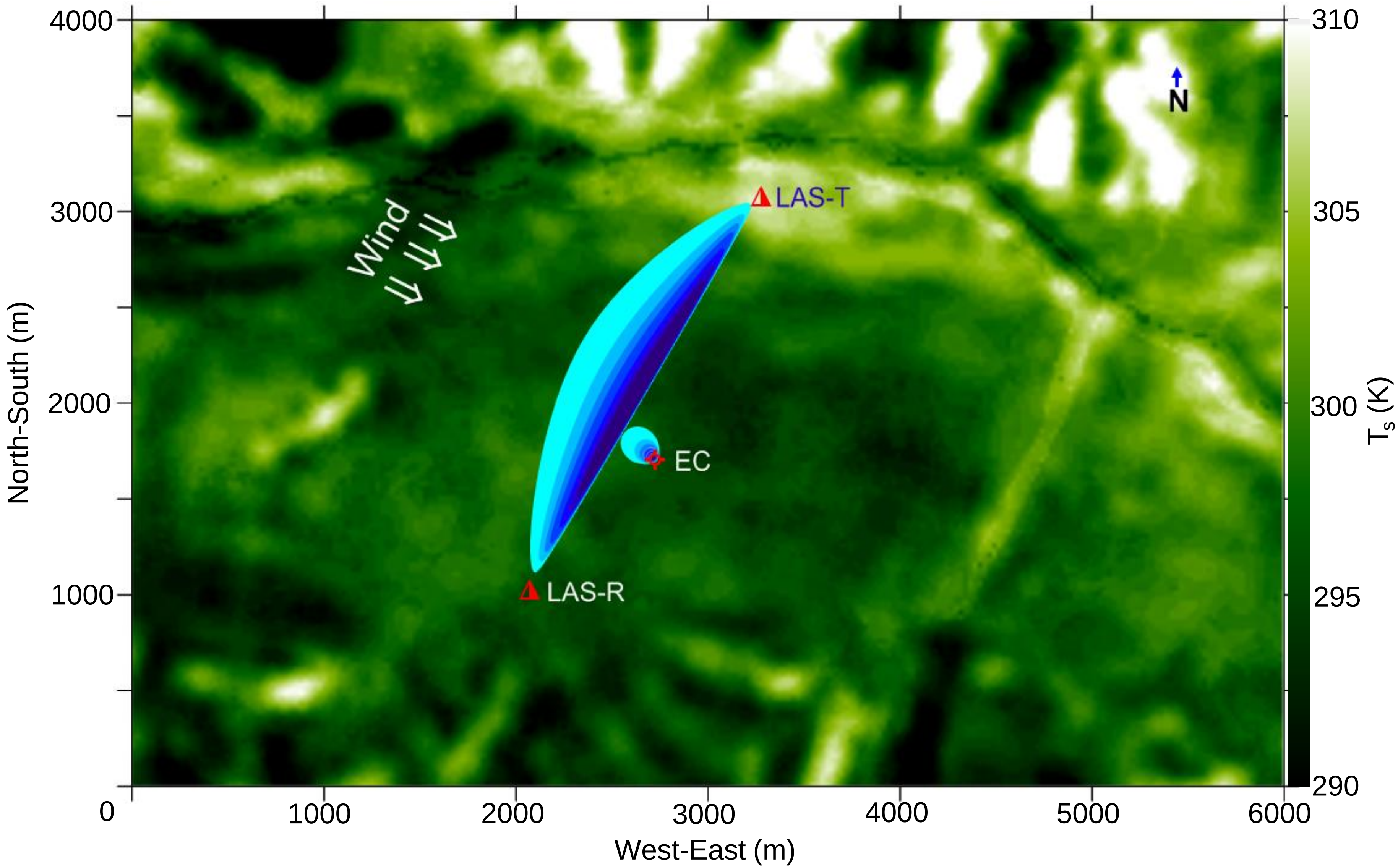
'Spectra' & 'Time-delay' methods are used to estimate C_T^2

The new fitting is drawn together with popular functions used.

New fitting is still in checking stage!







20080707-11:45 Background：Ts (from TM6); Observation: EC (3.15m), LAS (9.5m).
Wind dir., 311.6°. Wind speed, 0.77m/s (EC), 1.15m/s(LAS). u*=0.1778，H=97.55，z/L=-0.991
Footprint (outer to inner): 95%，90%，85%，80%，70%，60%，50% of the source area.

Conclusions (1/2)

- ‘Good data + Good modeling’ results ‘Good analysis & prediction’.
- Advanced instruments do not guarantee a data set of high quality. Specific knowledge and proper post-field data processing, with careful quality control and assessment, are essential.
- Approaching the Ground Truth is a goal of endless seeking. Our task is to know the uncertainties of the data from different observations, and improve the data quality step by step.

Conclusions (2/2)

- The SEB closure problem is now much clear. If a recommended turbulent data procedure of the EC system is followed, and, other components that contribute to the surface energy budget, esp. the soil heat storage in the upper layer, are properly included, the closure ratio can be up to 90% or higher.
- Single point EC system is limited in catching up the flux contributions from larger turbulent structures. LAS has the advantage of larger area coverage, however, it is limited by the using of MOST, which is semi-empirical.
- Source area (or footprint) analysis is necessary for surface parameter & flux studies. It is also essential for the validation and improvement of remote sensing algorithms and land surface models.

Acknowledgement

This study is supported by

- ***Project WATER***
- ***Project CEOP-AEGIS / FP7***

Thank You !