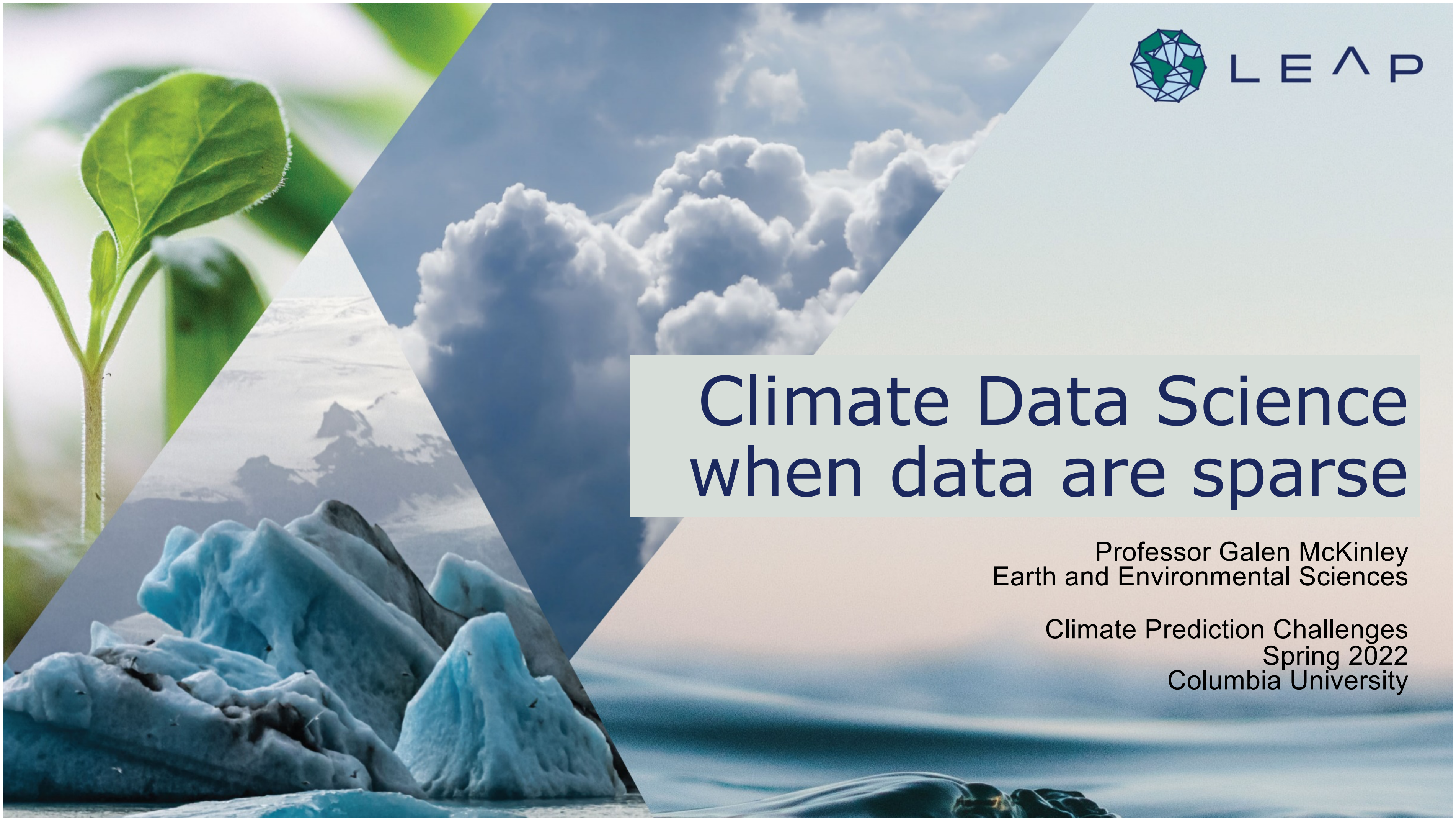


Climate Data Science when data are sparse

Professor Galen McKinley
Earth and Environmental Sciences

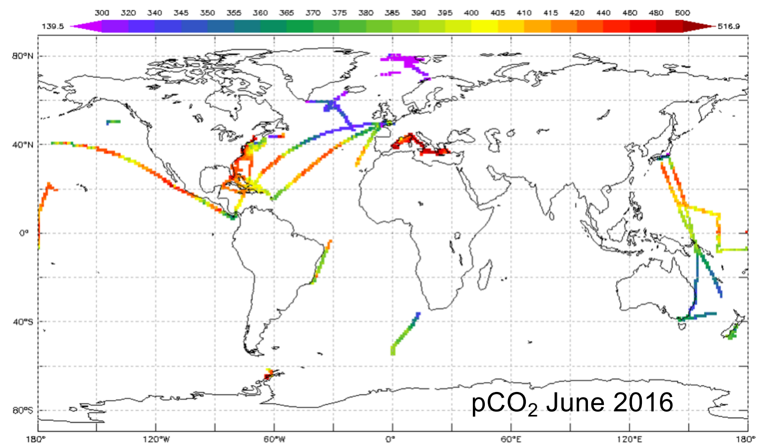
Climate Prediction Challenges
Spring 2022
Columbia University

Understanding Earth system processes often complicated by sparse data

- **Question: physical and ecological dynamics of lakes**
- **Data: a central buoy in only a few lakes**
- **PROJECT 2**

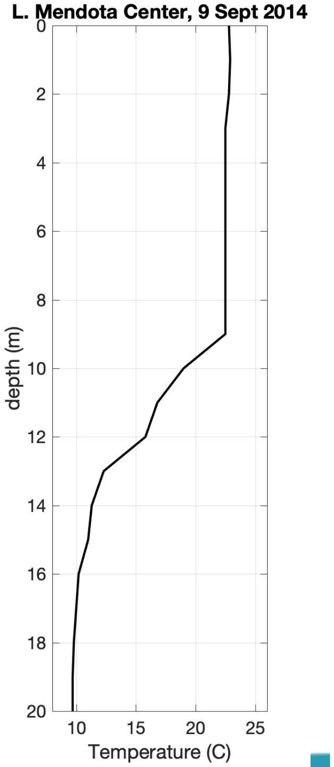
Lake Mendota, Madison, WI

- Question: processes of variable ocean CO_2 fluxes
- Data: infrequent ship transects
- PROJECT 3



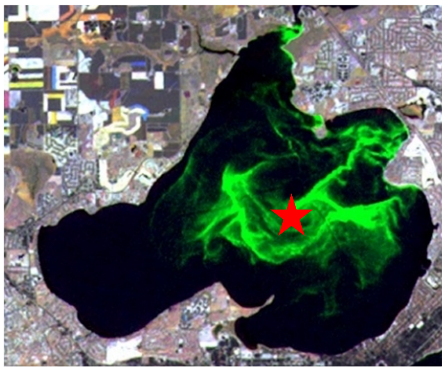
Observing Lake Mendota, Madison WI

Lake Mendota, Madison, WI

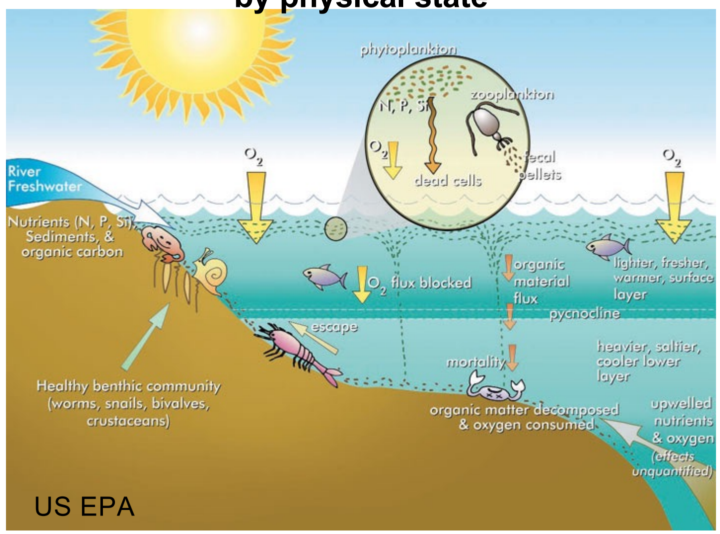


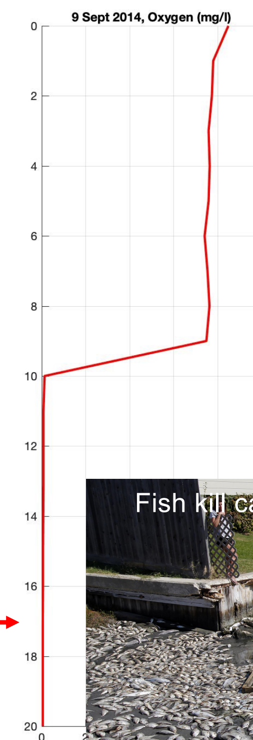
Why do we care about lake temperatures?

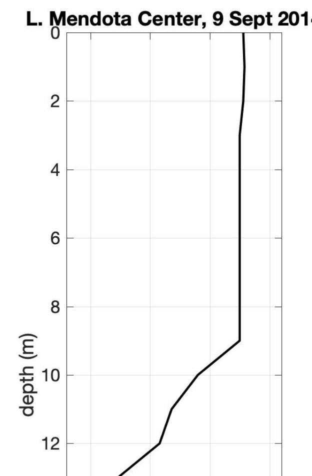
Lake Mendota, Madison, WI

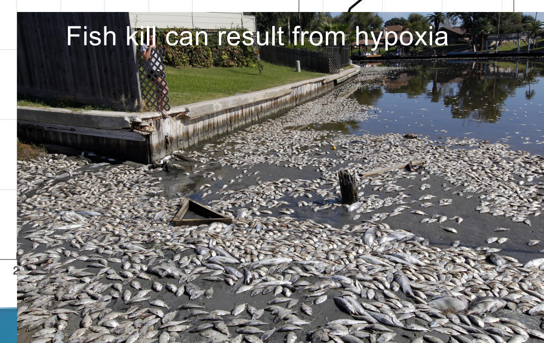


Lake Ecosystems strongly impacted by physical state





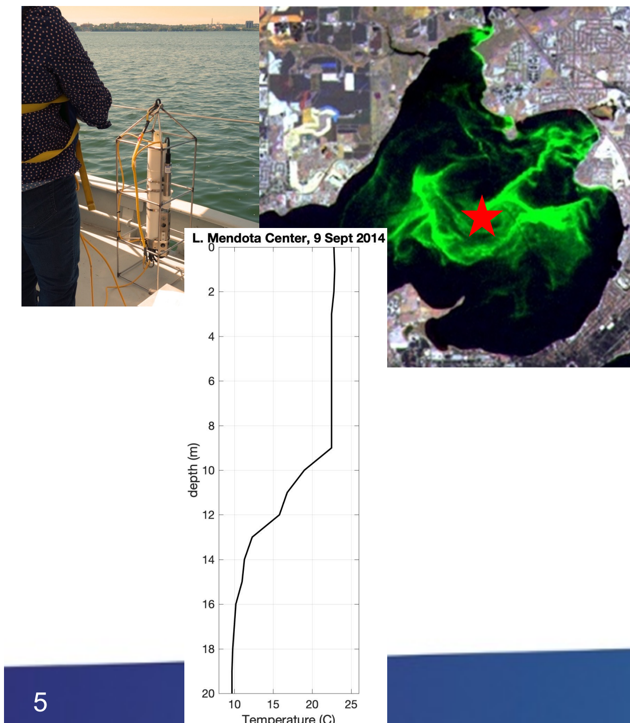


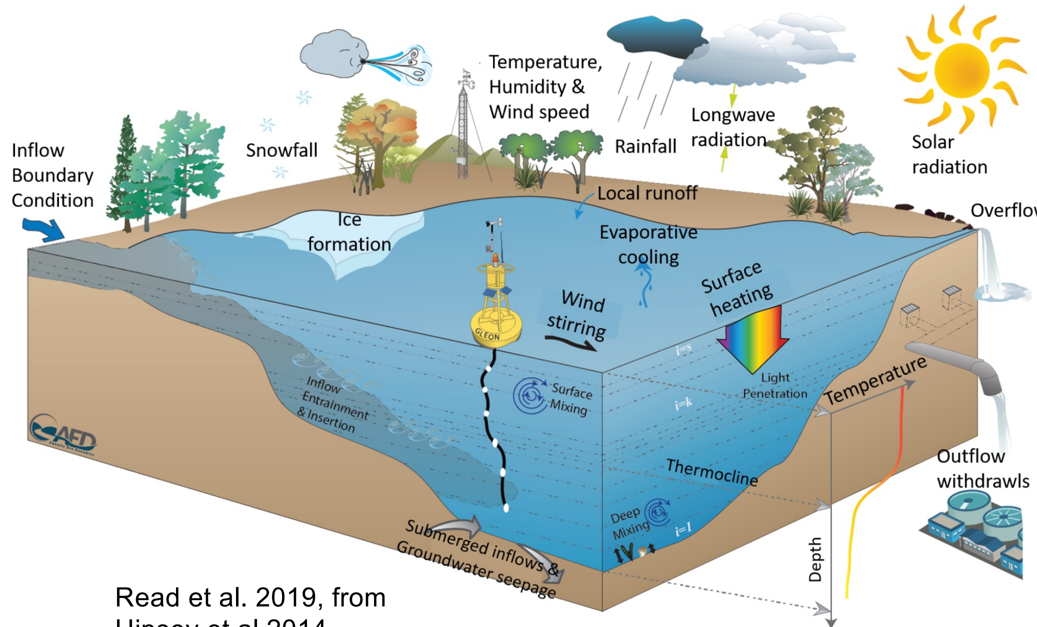


- Stratification encourages low O_2 at depth, reduces habitability for fish
- Humans make this problem worse by adding excess nutrients to lakes (eutrophication).
- Excess nutrients and warmer surface encourages harmful algal bloom (image at left)
- Surface temperatures increase and the length of the stratified season grows with climate change (O'Reilly et al. 2015 GRL, Sharma 2014)

Observation are sparse and labor-intensive, but we do know a lot about the physics

Lake Mendota, Madison, WI

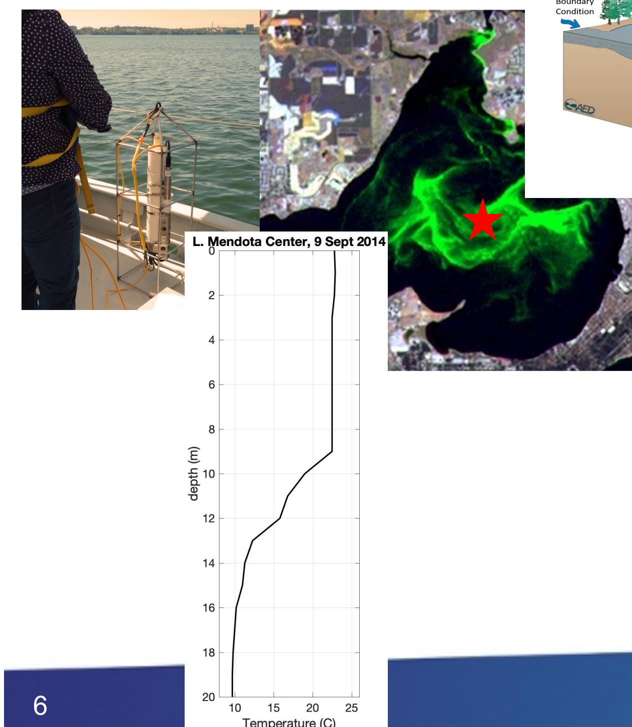


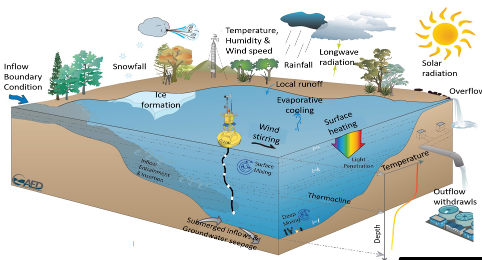


Read et al. 2019, from Hipsey et al 2014

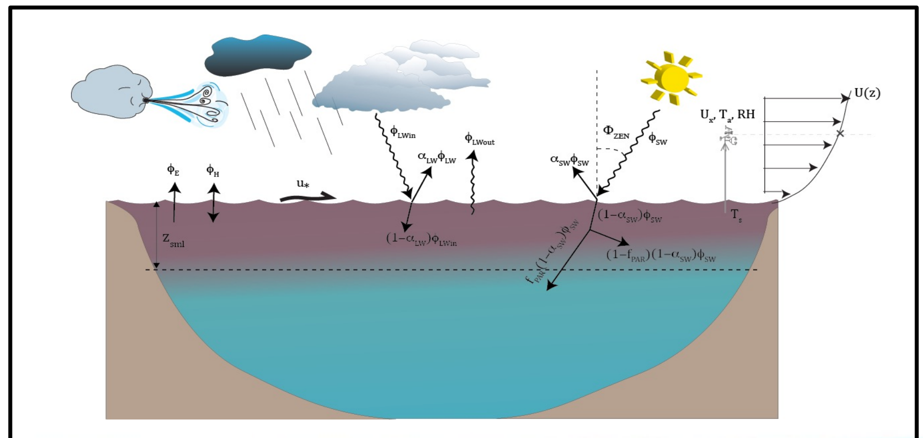
These physics can be embodied in a process model

Lake Mendota, Madison, WI



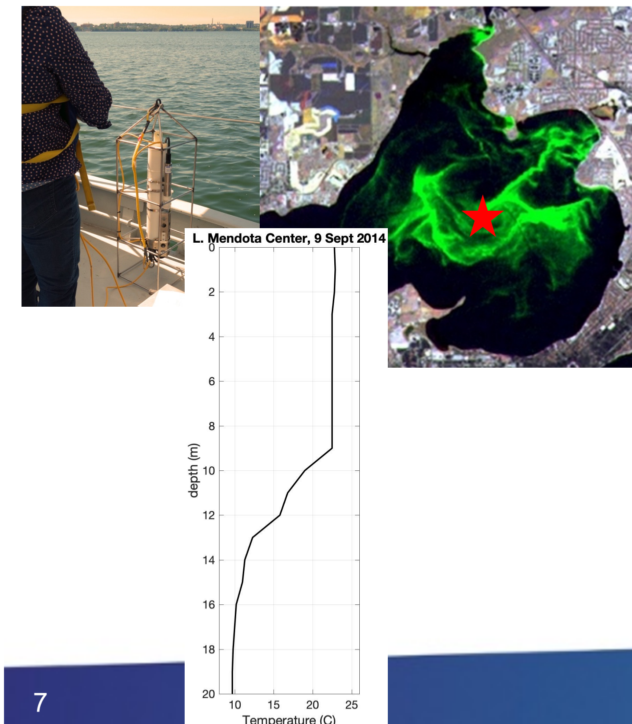


Process Model for 1D lake hydrodynamics

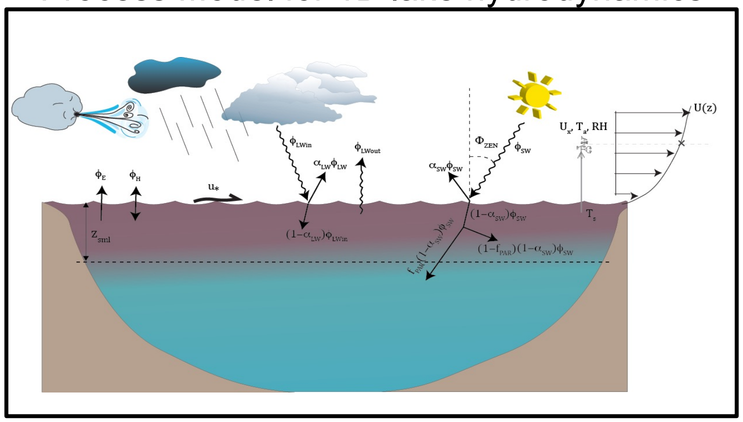


Process model to predict lake 1D temperatures

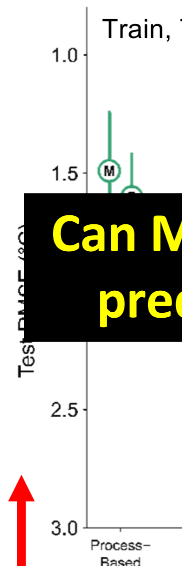
Lake Mendota, Madison, WI



Process Model for 1D lake hydrodynamics



Results



Can ML improve predictions?

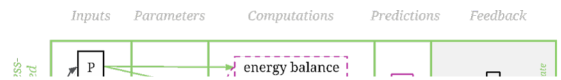
Better fit to independent data

Read, **Kumar** et al. 2019, *Water Resour. Res.*

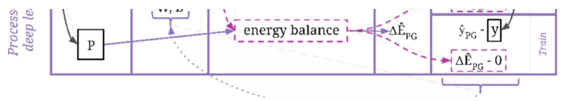
LEAP

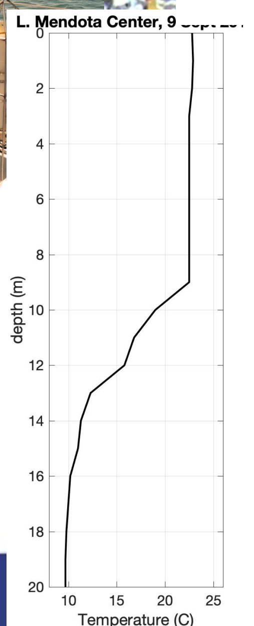
Physics-Guided ML blends process knowledge and deep learning for more robust predictions

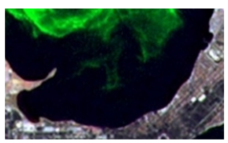
Lake Mendota, Madison,



(1) use deep learning / LSTM alone
(2) add pretraining with process model
+ energy conservation constraint



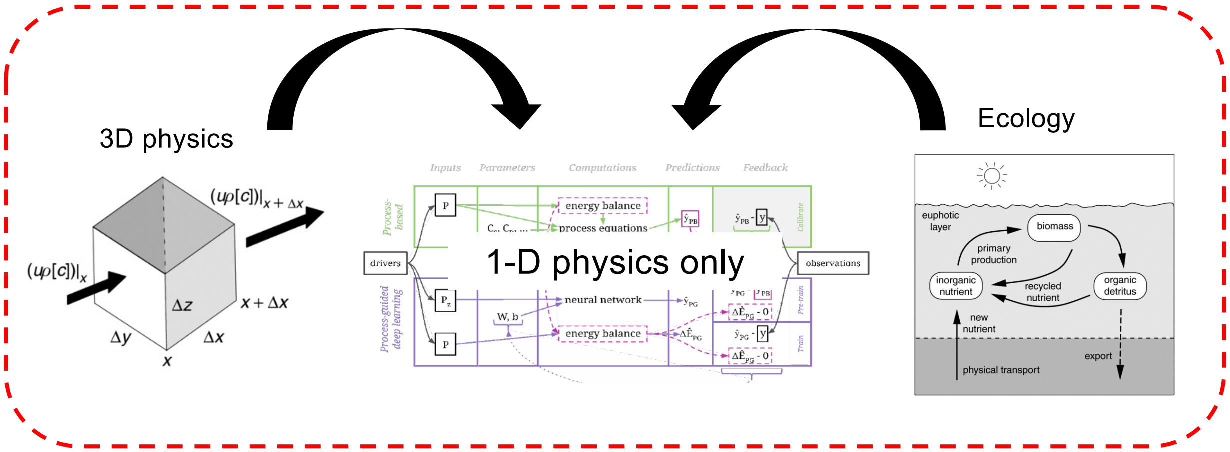




Read, **Kumar** et al. 2019, *Water Resour. Res.*

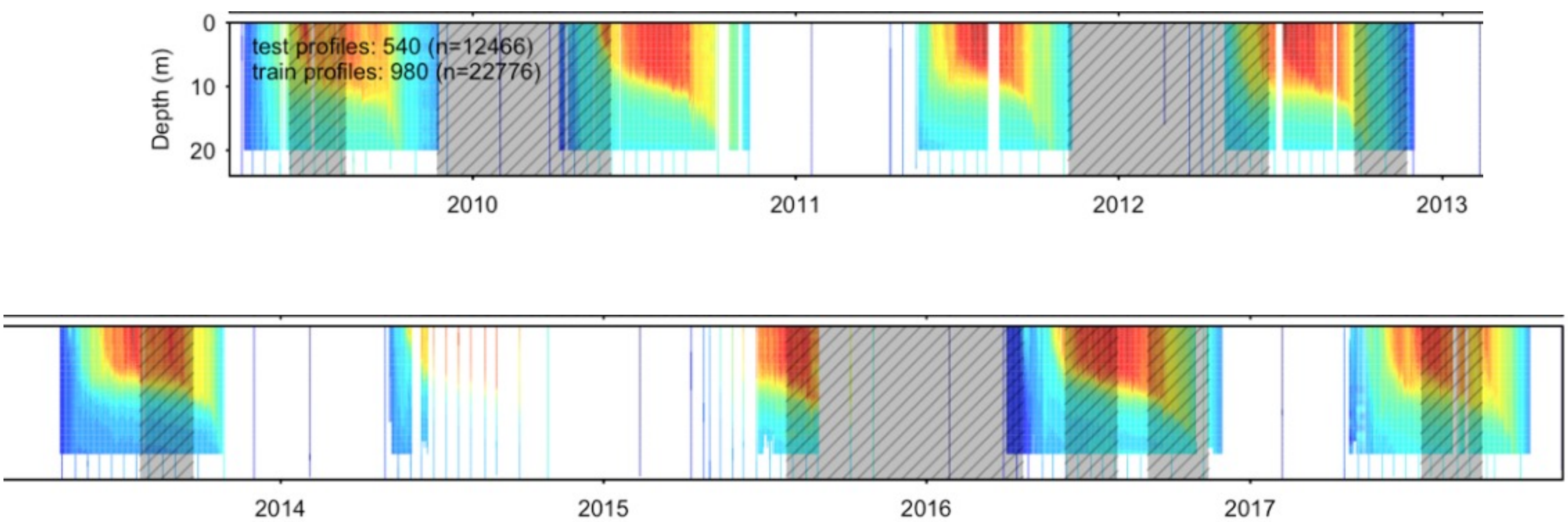
Next steps: estimate temperatures for 1000's of unobserved lakes, from surface data only

Next steps: extend to multiple interacting processes and spatial scales; other domains

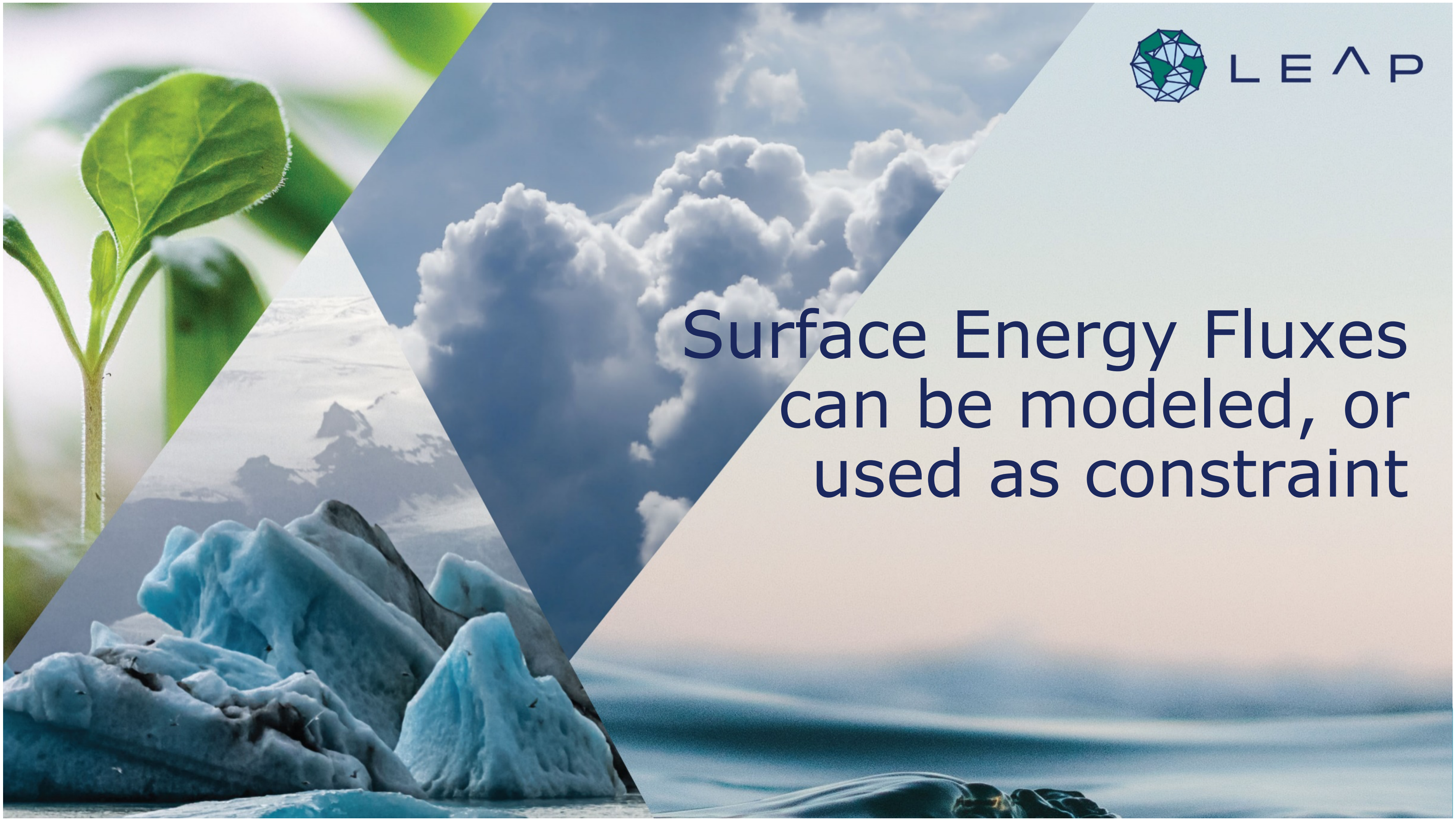



LEAP
Data for Read et al. 2019
See Supporting information at
https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019WR024922
(Scroll to bottom of page to download)

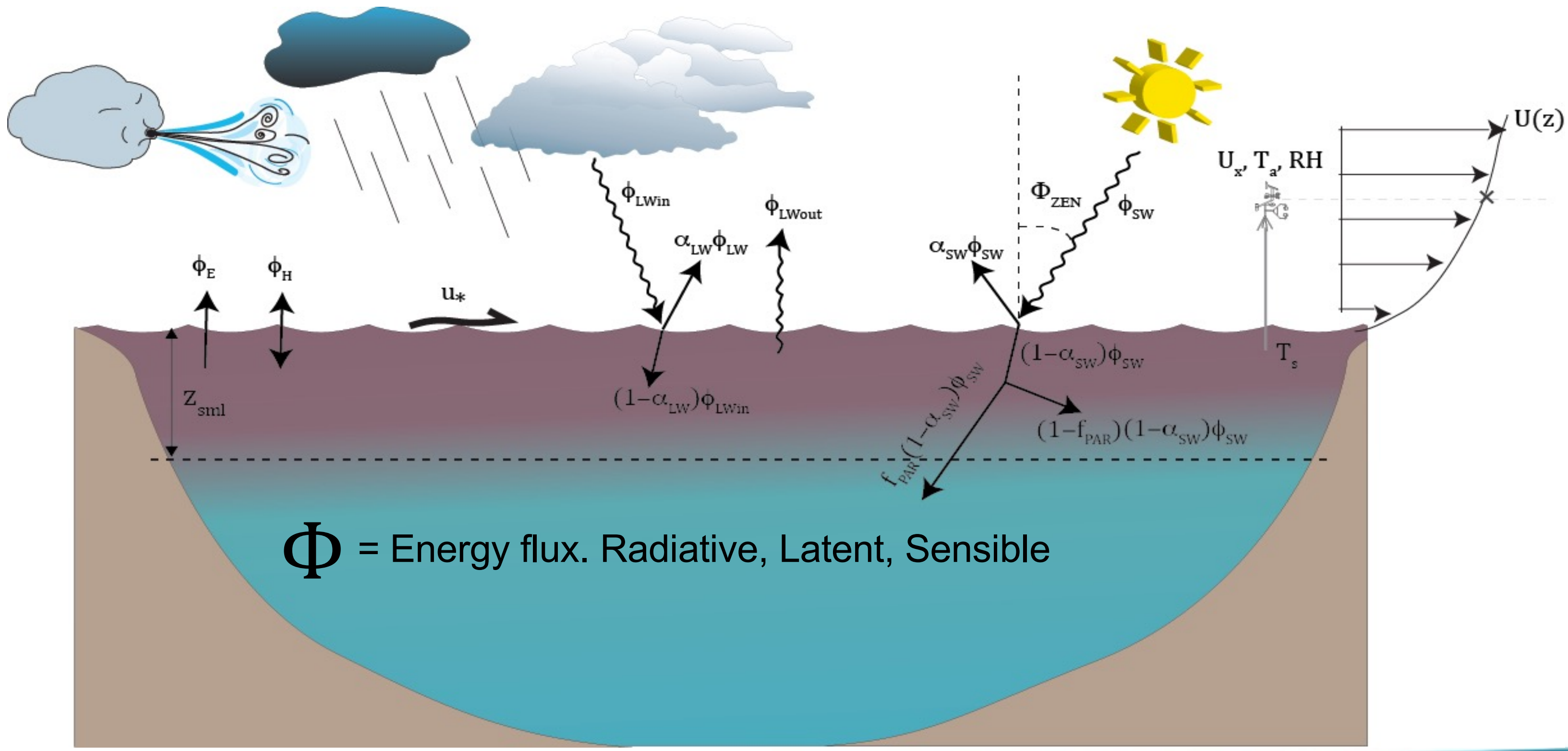
Lake Mendota



Test data covered with transparent gray

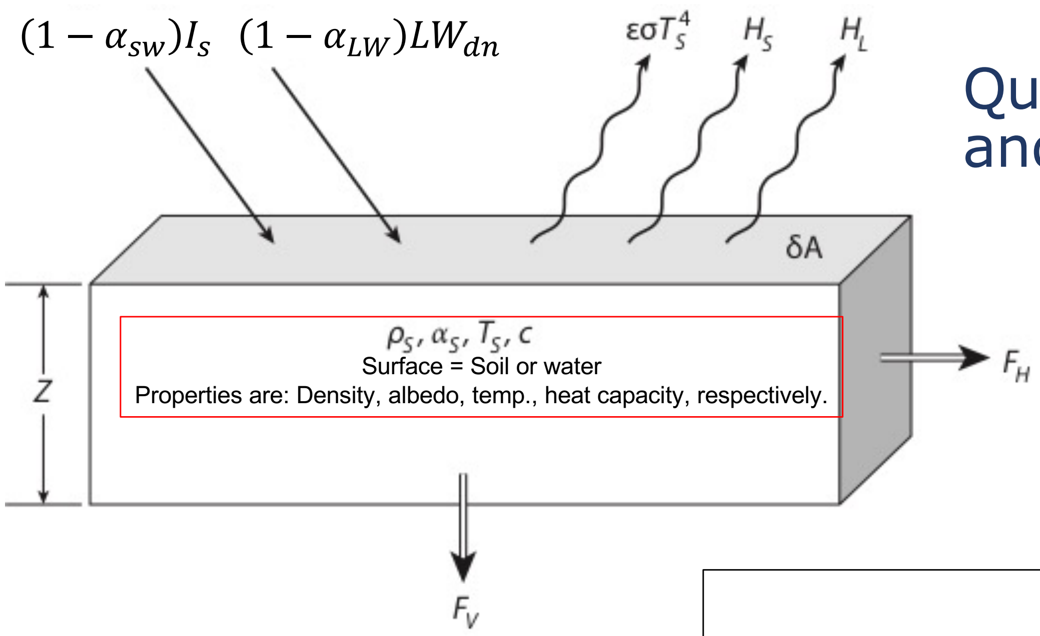

Surface Energy Fluxes
can be modeled, or
used as constraint

Lake model schematic



LEAP

Quantifying energy fluxes into and out of surface



Shortwave from Sun = Is = **R_sw_arr**
Albedo = α = **alpha_sw**, **alpha_lw**
Longwave from atmosphere = LW_{down} = LW_{dn} = **R_lw_arr**
Longwave from surface = $\varepsilon\sigma T_s^4$ = Lw_{up} = **R_lw_out_arr**
H_S = Sensible Heat = SH = **H**
H_L = Latent Heat = LE = LH = **E**
F_v, F_H = advection vertical, horizontal

Cook Fig 5.3

Water can warm or cool, but this change is constrained by the energy fluxes

$I_S(1-\alpha_{SW}) + LW_{dn}(1-\alpha_{LW}) - \varepsilon\sigma T^4 - H - E - F_v - F_H = C_w\, dT/dt$

where C_w is specific heat capacity of water

Fv, FH = 0 for a 1 dimensional column

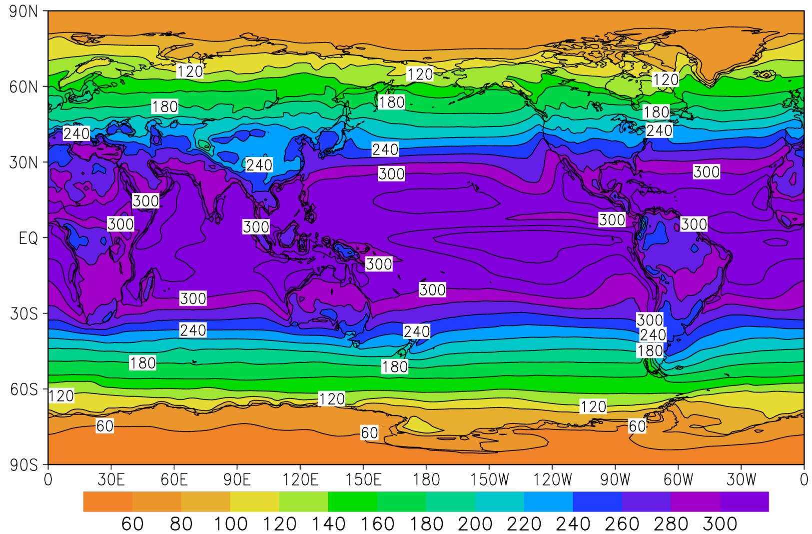
Thus, $I_S(1-\alpha_{SW}) + LW_{dn}(1-\alpha_{LW}) - \varepsilon\sigma T^4 - H - E = C_w\, dT/dt$

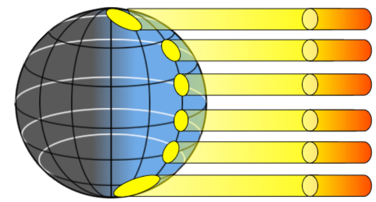
In Read et al. 2019, energy fluxes at surface and the integrated energy change ($C_w dT/dt$) must remain consistent

LEAP

Absorbed Solar Radiation
= Incident radiation*(1-albedo)

$$(1-\alpha)I_s$$

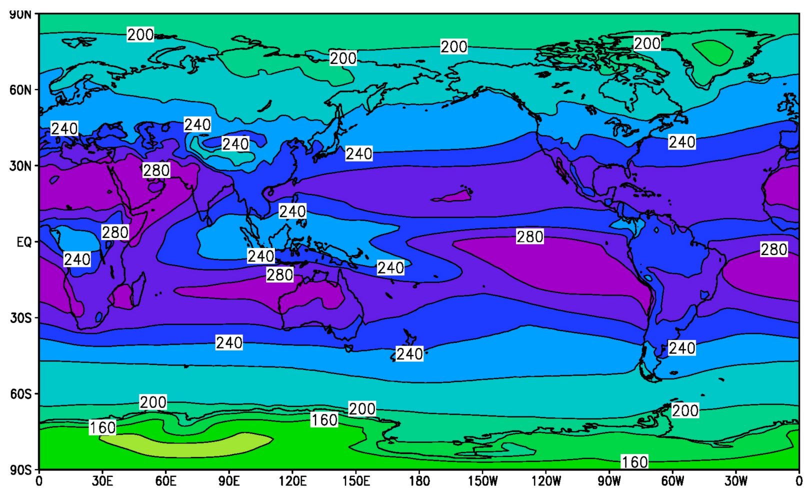




W/m^2, both panels

Outgoing Longwave
Radiation (OLR)

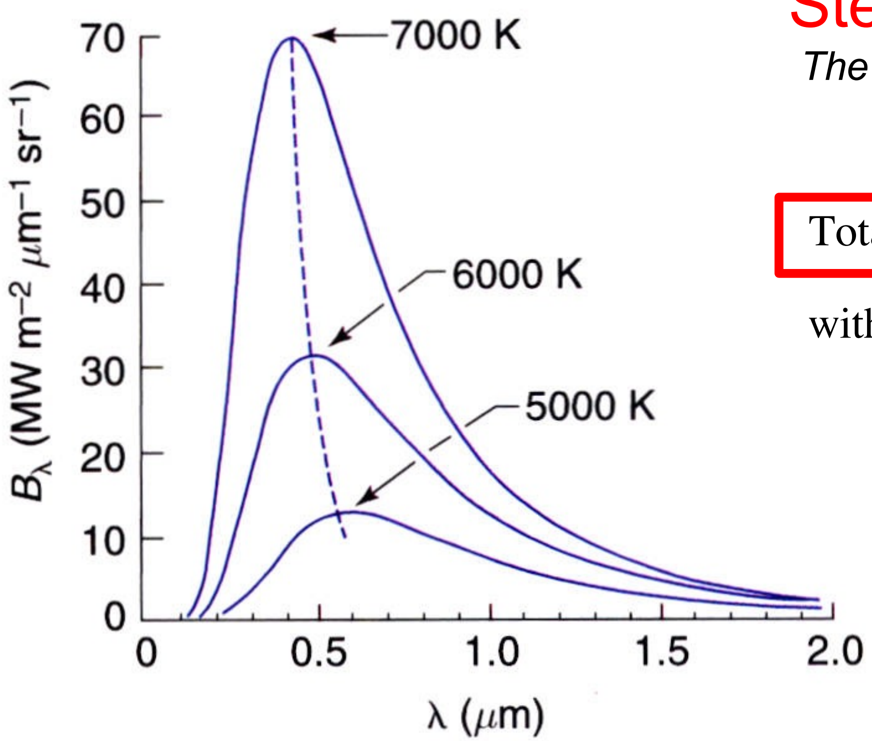
$$\varepsilon\sigma T^4$$



Cook Fig 5.5

Cook Fig 5.8

Longwave radiation calculated with Stefan-Boltzmann



Stefan – Boltzmann Law

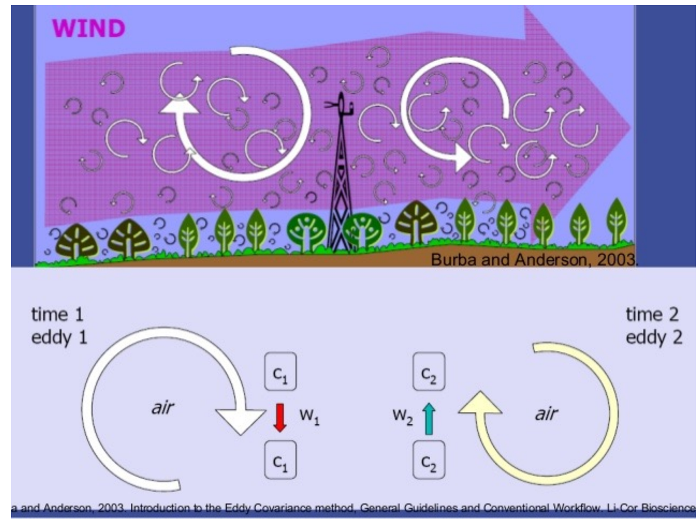
The integral across all wavelengths of the previous two

$$\boxed{\text{Total Energy Emitted (W/m}^2\text{)} = \sigma T^4}$$

$$\text{with } \sigma = 5.67x10^{-8} \frac{W}{m^2 K^4}$$

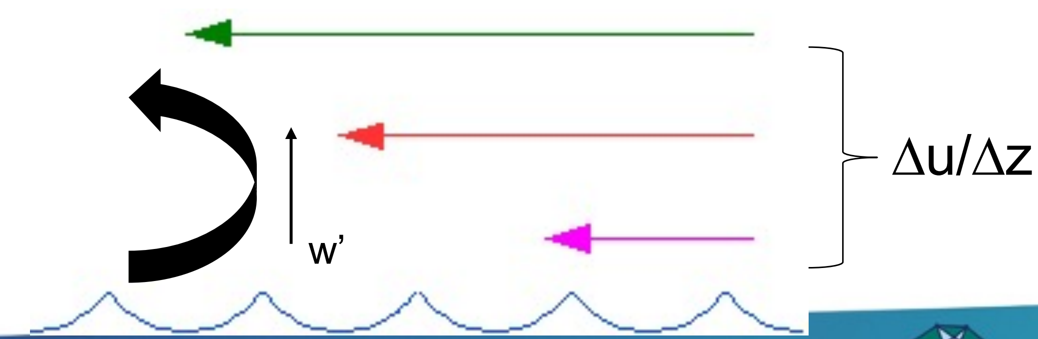
Parameterizing SH and LE as surface flux dependent on atmospheric turbulence

- From physical reasoning, and checking units, we can reason that
- $SH = c_p \rho w T$
    - c_p = specific heat of air; ρ = air density (kg/m^3)
- $LE = L \rho w q$
    - L = latent heat of vaporization
- But we need to be a little more specific about the time and space scales for w (m/s), T(K) and q (g $H_2 0$/kg) …



Key assumption in first-order efforts to estimate turbulence

- Vertical motions driven turbulently (w') will be proportional to the vertical gradient of horizontal wind speed
- i.e. if winds aloft are slow, turbulence will be damped. If they are fast, there will be more turbulence



Parameterizing SH and LE

- Then
  - $SH = c_p \, \rho \, \overline{(w' T')}$
  - $LE = L \, \rho \, \overline{(w' q')}$
- Based on the reasoning of previous slide :

  $\overline{(w' T')} \sim U_R \nabla_z T$

  $\overline{(w' q')} \sim U_R \nabla_z q$
- Then discretize to two layers (surface and z_r)
  - $\nabla_z T = T_s - T_a(z_r)$
  - $\nabla_z q = q_s - q_a(z_r)$
- Then empirically determine C_{DH} and C_{DE}
  - $C_D = C_{DH} = C_{DE} = 0.0013$ to 0.002

Full SH and LE equations

$$SH = c_p \rho C_{DH} U_r (T_s - T_a(z_r))$$

$$LE = L \rho C_{DE} U_r (q_s - q_a(z_r))$$

Water can warm or cool, but this change is constrained by the energy fluxes

$I_S(1-\alpha_{SW}) + LW_{dn}(1-\alpha_{LW}) - \varepsilon\sigma T^4 - H - E - F_v - F_H = C_w\, dT/dt$

where C_w is specific heat capacity of water

Fv, FH = 0 for a 1 dimensional column

Thus, $I_S(1-\alpha_{SW}) + LW_{dn}(1-\alpha_{LW}) - \varepsilon\sigma T^4 - H - E = C_w\, dT/dt$

In Read et al. 2019, energy fluxes at surface and the integrated energy change ($C_w dT/dt$) must remain consistent