Hydrological Modelling in a data rich era

Some idea about modelling

Takashi Murakami

Rigon R.

Warredoc, Perugia January 30, 2019

Aristotle said a bunch of stuff that was wrong. Galileo and Newton fixed things up. Then Einstein broke everything again. Now, we've basically got it all worked out, except for small stuff big stuff, hot stuff, cold stuff, fast stuff, heavy stuff, dark staff, ***turbulence,*** *and the concept of time.*

Hydrological models**, like any Physics**

requires data

Physics is based on observations (actually well designed experiments)

"In general, we look for a new law by the following process. ***First, we guess it*** *(audience laughter), no, don't laugh, that's really true. Then we compute the consequences of the guess, to see what, if this is right, if this law we guess is right, to see what it would imply and* ***then we compare the computation results to nature****, or we say compare to experiment or experience, compare it directly with observations to see if it works.*

If it disagrees with experiment, it's wrong. *In that simple statement is the key to science. It doesn't make any difference how beautiful your guess is, it doesn't matter how smart you are who made the guess, or what his name is… If it disagrees with experiment, it's wrong. That's all there is to it."*

R. Feyman

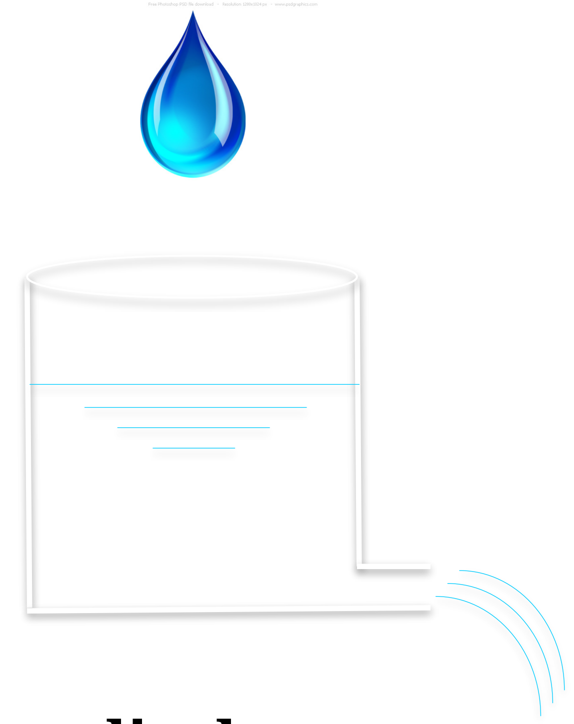
https://youtu.be/OL6-x0modwY

Feynman was to good not to cite it, but he referred to laws, which is exactly not the same as models. For models maybe, **the** question is:

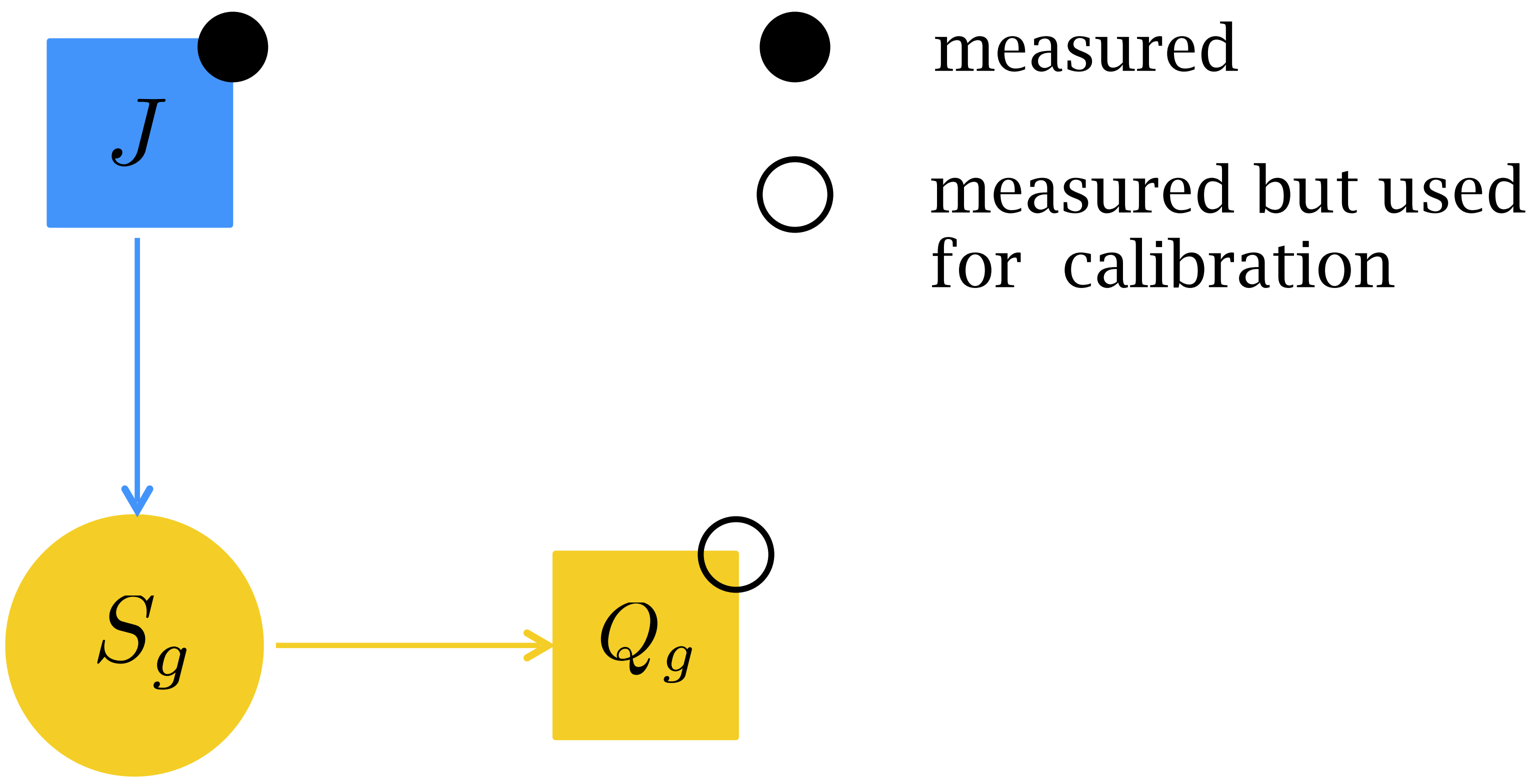
Where do datasets can be integrated best ?

The answer, obviously , depends on models construction, as we ‘ll try to clarify.

Once upon a time there were **simple** hydrological models



Emphasis was on predicting **discharges**



$$\frac{dS_g}{dt} = J^{\bullet}(t) - Q_g^{\circ}(t)$$

$$S_g(t) = kQ_g(t)$$

The linear model

Everything is known of the linear model. In particular

$$Q(t) = \int_{t_0}^{t} e^{-(t-t_{in})/k} J(t_{in}) dt_{in} \qquad *$$

and

$$E[\ \underbrace{t - t_{in}}_{\text{Travel time **}}\ ] = k$$

* This formula is correct but a little misleading because it move the attention from the budget to the discharges only

** Is a little more complicate than that. See Botter et al., 2011 and Rigon et al., 2016

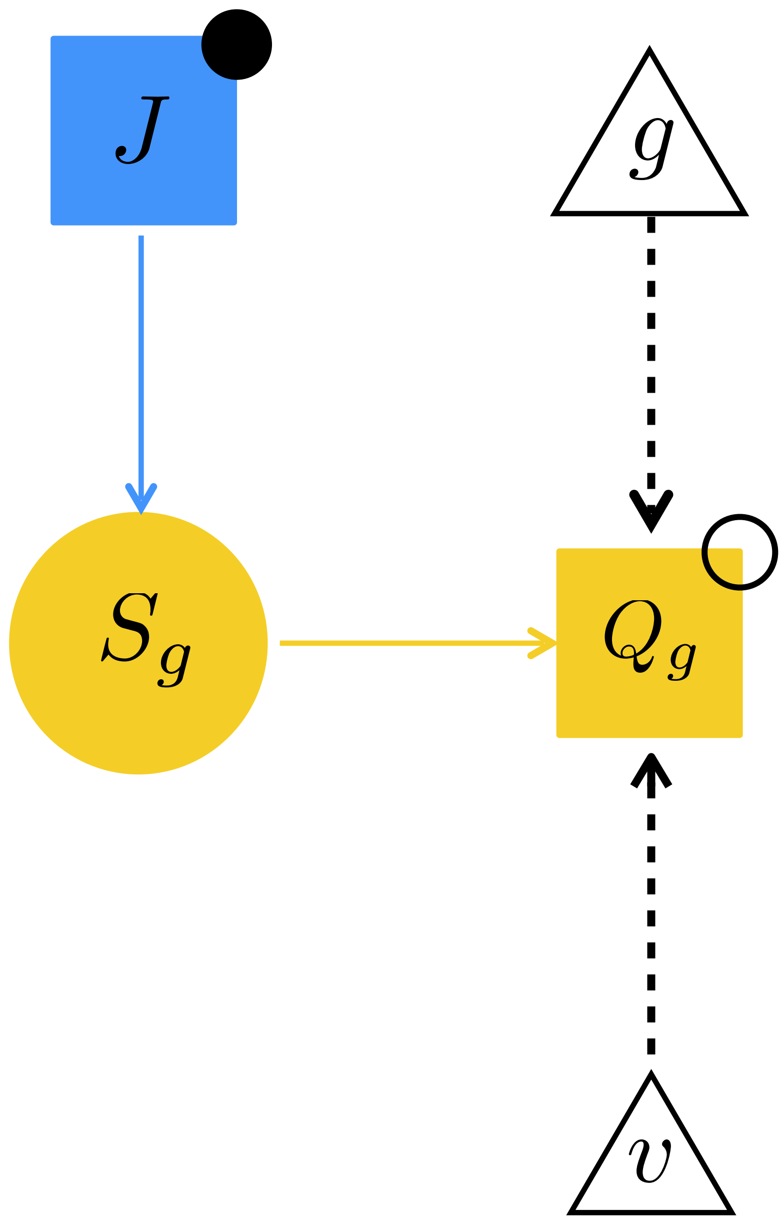
Does it know about geology ?

Does it know about vegetation ?

NO !

So any geological and vegetation measure is useless

Sam Cook, What a wonderful world
….
But I do know that I love you
And I know that if you love me, too
What a wonderful world this would be



Unless $k = k(g, v)$ *v*egetation and *g*eology are **controllers** of k

Because measures become important, either state variables or parameters must have a hook to measures.
Not necessarily measures have to be the parameters or the state variables. But "functions" of it.

i.e. $k = k(g, v)$

Measure are useless without a model to give them significance

Someone says, "Well, data can do it all alone."

In fact if

$$\frac{dS_g}{dt} = J^\bullet(t) - Q_g^\bullet(t)$$

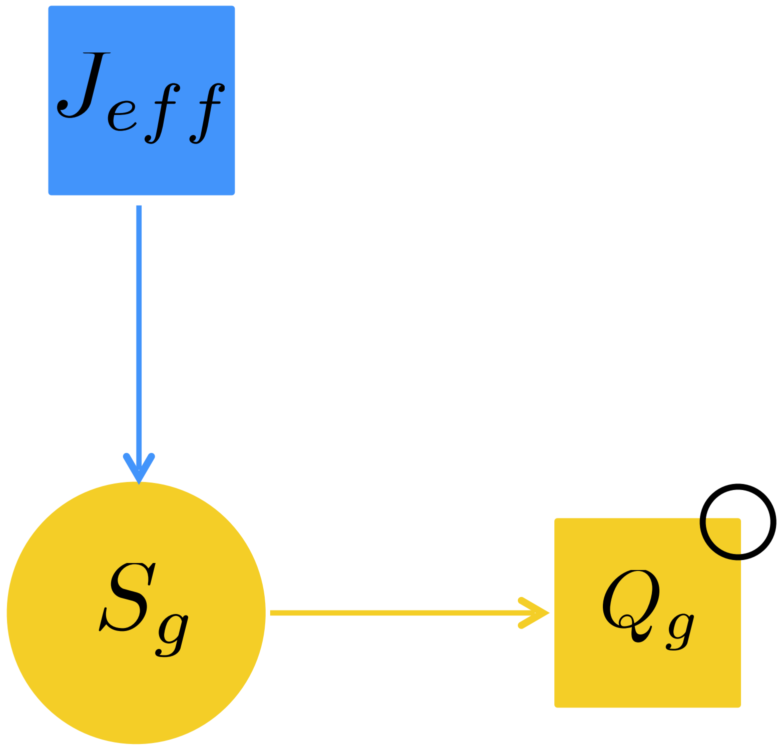
the storage variation is completely known (as soon as we have measures)

If we do it we will observe that

$$\frac{dS_g}{dt} > 0$$

so the model is wrong because water storage does not increase indefinitely. **Models are useless without measures.**

The model is not enough. This the classical answer to the problem



J_{eff} is that part of

precipitation for which

$$\overline{\frac{dS_g}{dt}} \approx 0$$

We lived with it for decades but it has quite a few problems

We have to introduce a lot of hypothesis about what J_{eff} is.

We can see the morphology of the catchments*

VOL. 15, NO. 6 WATER RESOURCES RESEARCH DECEMBER 1979

The Geomorphologic Structure of Hydrologic Response

IGNACIO RODRÍGUEZ-ITURBE AND JUAN B. VALDÉS

Graduate Program in Hydrology and Water Resources, Universidad Simón Bolivar, Caracas, Venezuela

A unifying synthesis of the hydrologic response of a catchment to surface runoff is attempted by linking the instantaneous unit hydrograph (IUH) with the geomorphologic parameters of a basin. Equations of general character are derived which express the IUH as a function of Horton's numbers R_A, R_B, and R_L; an internal scale parameter L_Ω; and a mean velocity of streamflow v. The IUH is time varying in character both throughout the storm and for different storms. This variability is accounted for by the variability in the mean streamflow velocity. The underlying unity in the nature of the geomorphologic structure is thus carried over to the great variety of hydrologic responses that occur in nature. An approach is initiated to the problem of hydrologic similarity.

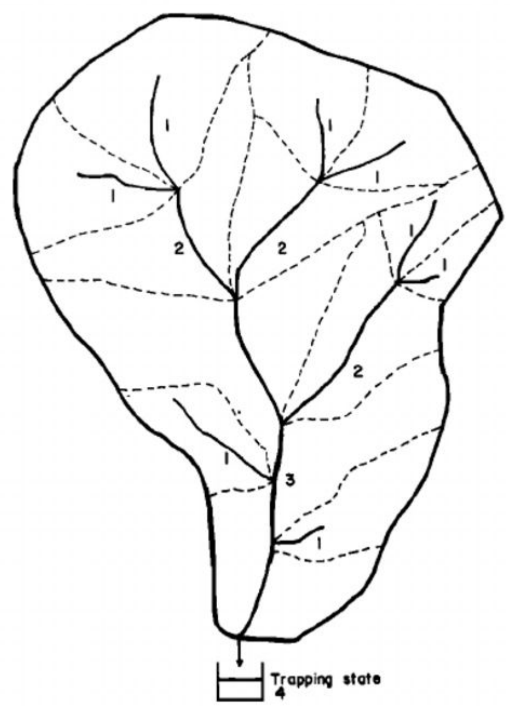


Fig. 1. Third-order basin with Strahler's ordering system and its trapping state.

*It was not that easy 40 years ago. No DEMs. No LIDAR at that time

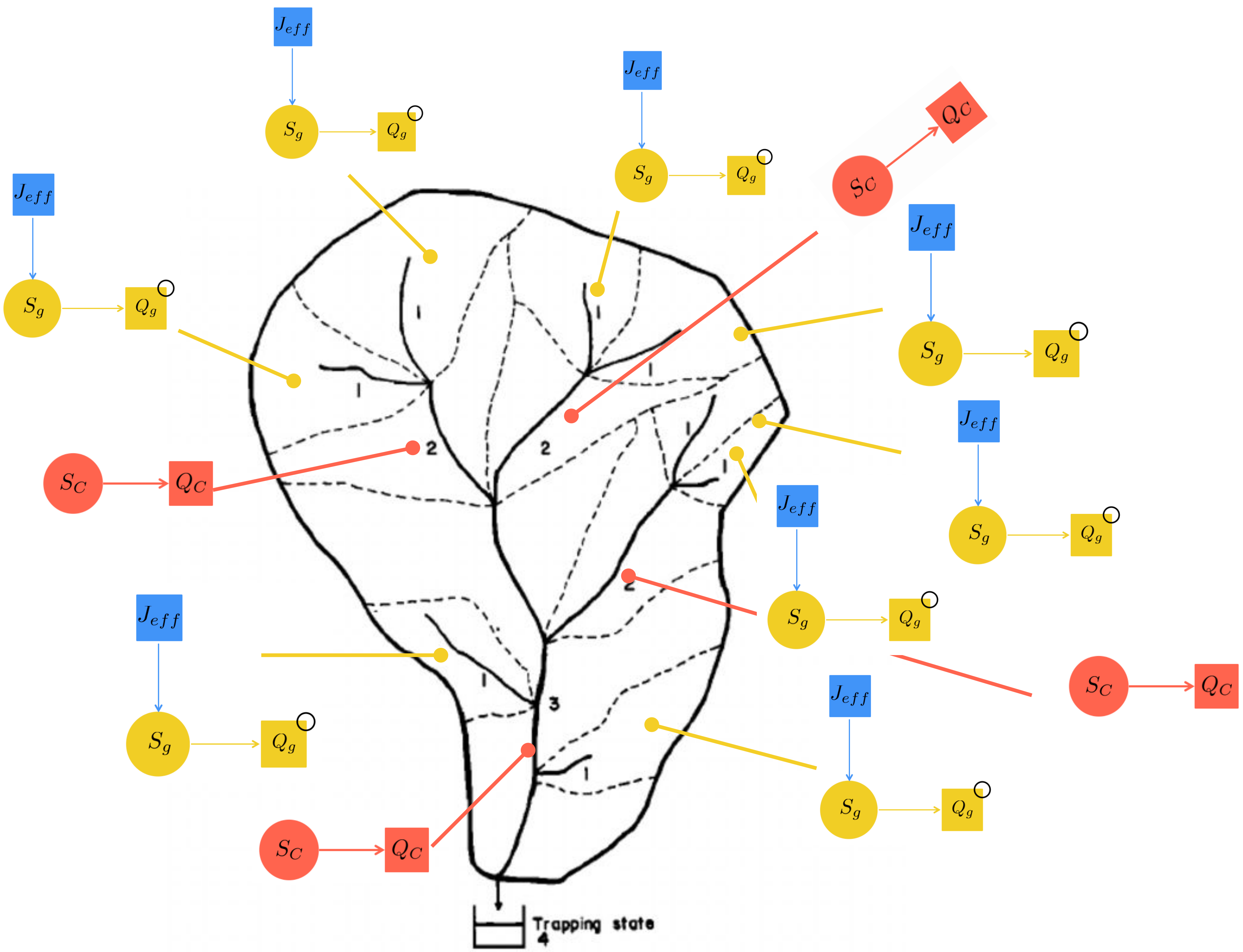
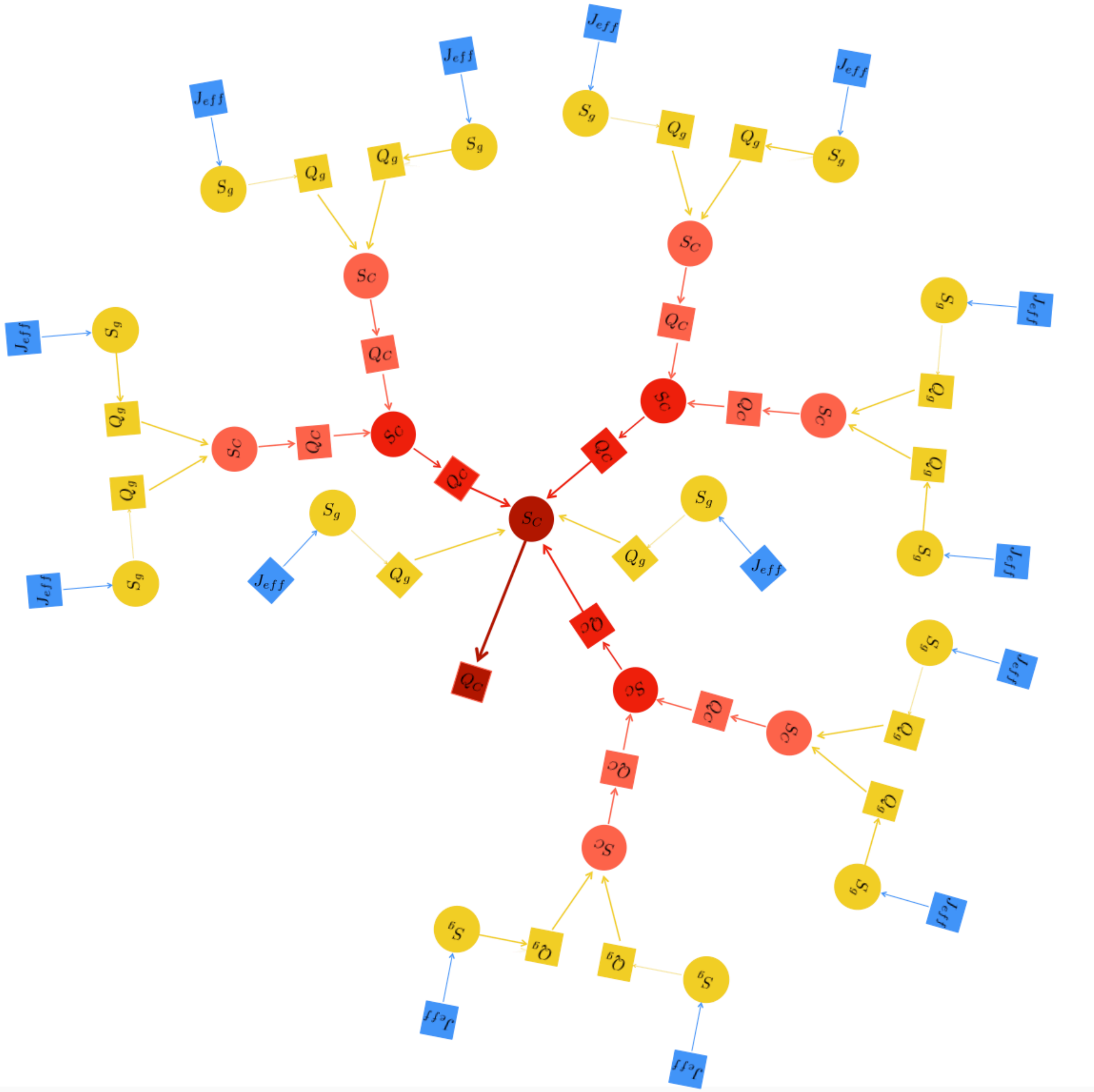


Fig. 1. Third-order basin with Strahler's ordering system and its trapping state.

R. Rigon



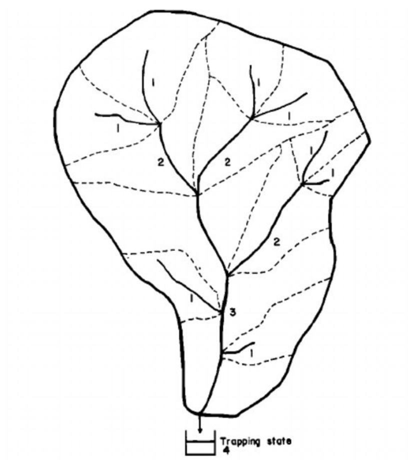


Fig. 1. Third-order basin with Strahler's ordering system and its trapping state.

In principle we can accomodate: spatially distributed rainfall and geomorphic characteristics (the latter trough k and contributing areas). The final discharge is the convolution of discharges along path. See Rigon et al., 2016a

Eventually people became able to measure# **travel times** with tracers

$$E[\underbrace{t - t_{in}}_{\text{Travel time}}] = k$$

And measures messed up everything

McDonnell JJ. 1990. A rationale for old water discharge through macropores in a steep humid catchment. Water Resources Research 26: 2821–2832.

again not so simple, indeed

Botter, G., Bertuzzo, E., & Rinaldo, A. (2010). Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox. Water Resources Research, 46(3). http://doi.org/10.1029/2009WR008371

fix it up distinguishing among uniform sampling and preferential sampling of water ages.
See also:

Benettin, P., Soulsby, C., Birkel, C., Tetzlaff, D., Botter, G., & Rinaldo, A. (2017). Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments. Water Resources Research, 53(3), 1864–1878. http://doi.org/10.1002/2016WR020117

Botter, G., Bertuzzo, E., & Rinaldo, A. (2011). Catchment residence and travel time distributions: The master equation. Geophysical Research Letters, 38(11).

Rigon, R., Bancheri, M., & Green, T. R. (2016). Age-ranked hydrological budgets and a travel time description of catchment hydrology. Hydrology and Earth System Sciences, 20(12), 4929–4947. http://doi.org/10.5194/hess-20-4929-2016

Rinaldo, A., Benettin, P., Harman, C. J., Hrachowitz, M., McGuire, K. J., van der Velde, Y., et al. (2015). Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes. Water Resources Research, 51(6), 4840–4847. http://doi.org/10.1002/2015WR017273

They say the J_{eff} has to do with infiltration. So various models of infiltration were used.

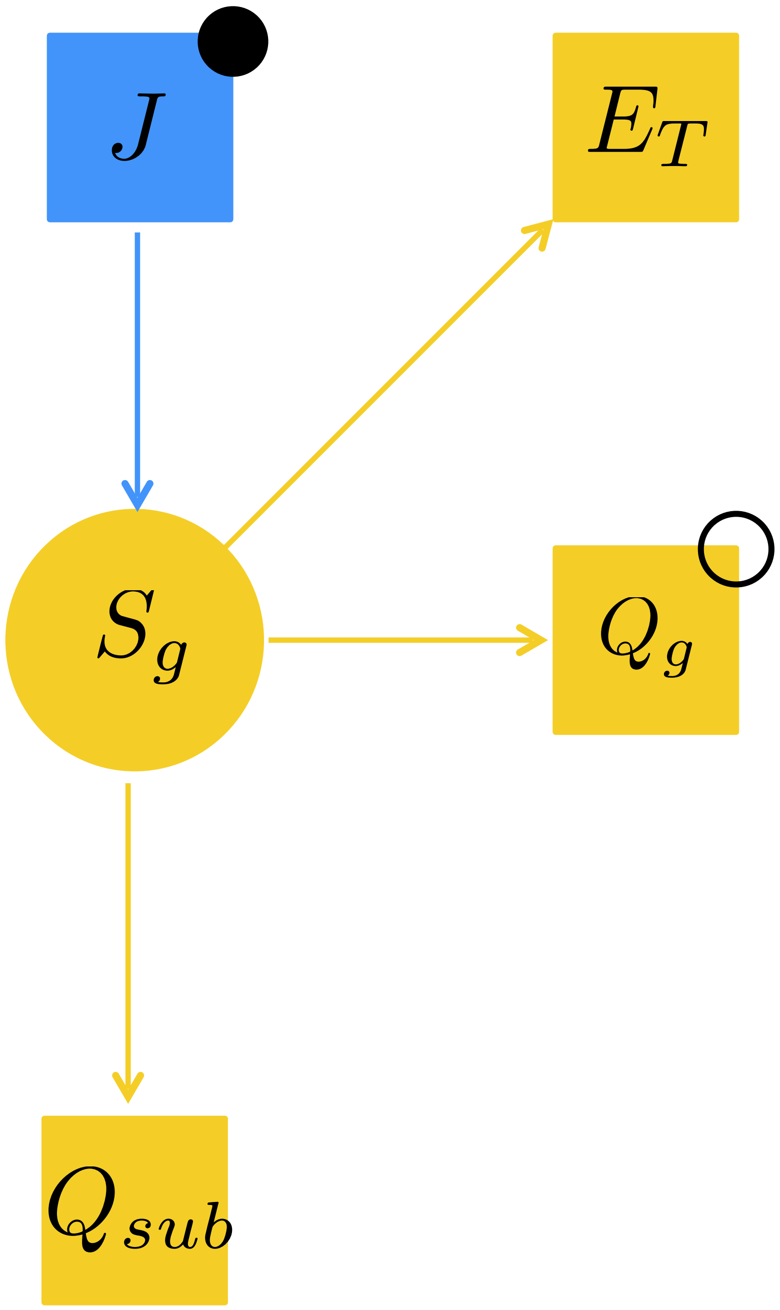
Obviously when you introduce a model for infiltration, you introduce **new variables and new parameters**.

Complication grows, you need more measures

Guercino- Sisifo, 1636

This model is better

$$\frac{dS_g}{dt} = J^\bullet(t) - Q_g^\bullet(t) - E_T(t) - Q_{sub}(t)$$



The problem now is undetermined, even if we assume

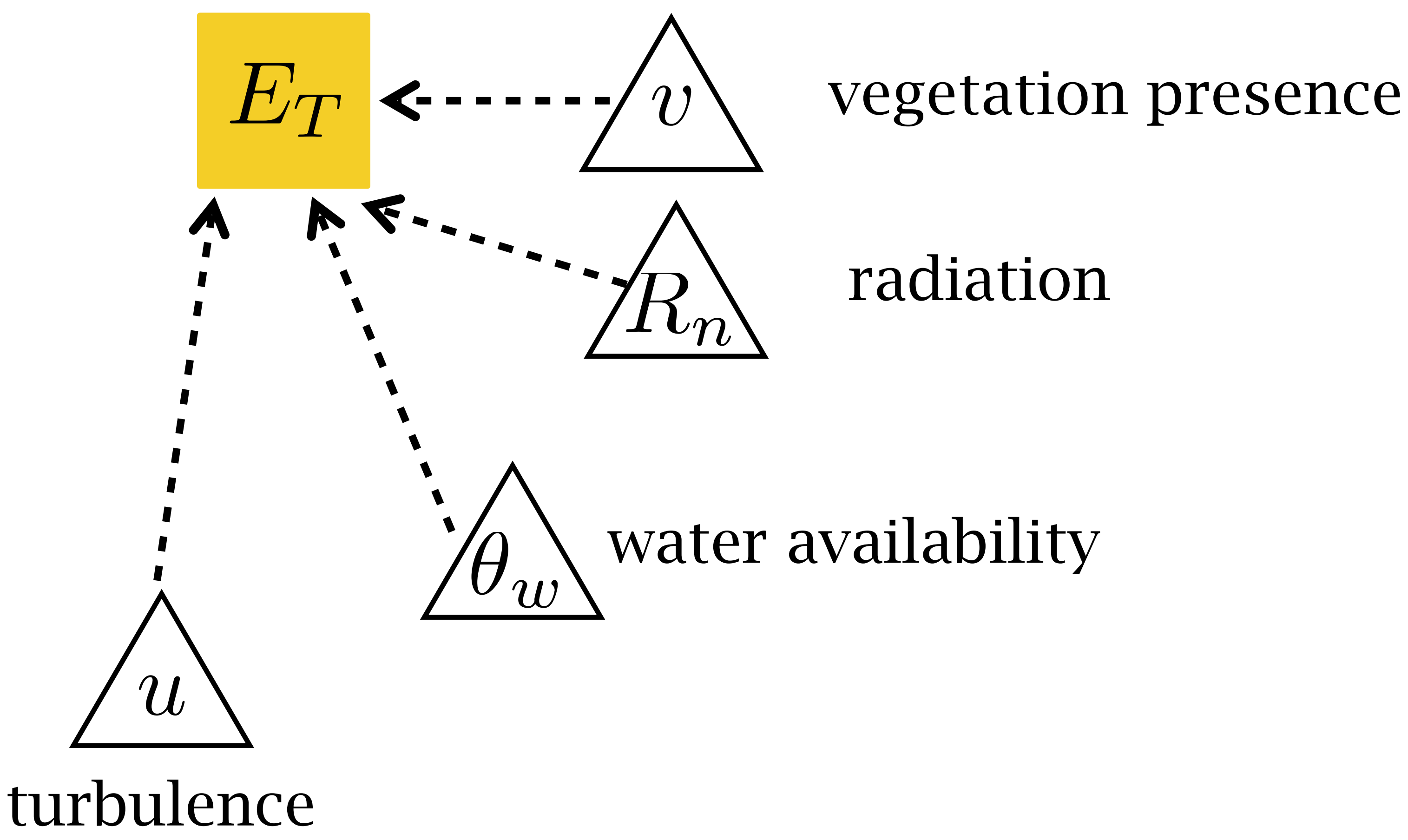
$$S_g(t) = kQ_g(t)$$

Can we measure

$$E_T(t), \quad Q_{sub}(t)$$

to close the budget ?

But what is evaporation function of ?



Evaporation and transpiration are from 30% to 60% of the whole hydrological budget, depending on climate

The traditional equation for E_T

$$E_T = \rho C_E \frac{\epsilon}{p} \overline{u} (e^*(z_0) - e(z)) \quad (1)$$

Dalton's law (turbulent transport of water vapor)

Symbol	Name	Unit
C_E	Dalton number	[-]
e	Partial pressure of water vapor	[F L^{-2}]
e^*	Partial pressure of water vapor at equilibrium	[F L^{-2}]
E_T	Evapotranspiration flux	[E T^{-1} L−2]
p	Air Pressure	[F L^{-2}]
$\overline{u}$	Horizontal direction component of mean velocity	[L T^{-1}]
ϵ	0.622	[-]
ρ	Dry air density	[M L^{-3}]

2 $$R_S = \lambda E_T + H + R_{ll}$$

Stationary energy budget (no temporal accumulation of energy in the control volume)

Symbol	Name	Unit
E_T	evapotranspiration flux	[M T^{-1} L^{-2}]
H	Sensible heat flux	[E T^{-1} L^{-2}]
R_{ll}	Upward longwave radiation	[E T^{-1} L^{-2}]
R_s	Shortwave radiation	[E T^{-1} L^{-2}]
λ	specific entalphy of vaporization	[E M^{-1}

$$H = a\,\rho C c_P \overline{u}\,(T(z_0) - T(z)) \qquad (3)$$

Turbulent transport of heat

$$a = \begin{cases} 1 & \text{if soil or water surface} \\ 2 & \text{if leaf} \end{cases}$$

Symbol	Name	Unit
C	Evaporation factor	[-]
c_p	Specific heat capacity	[L^2 M T^{-2} Θ^{-1}]
T	Tempeature	[Θ]

Because we hav four unknown and three equation, we need a further equation

4 $$e^*(z_0) - e(z) = \delta e(z) + \Delta(T(z_0) - T(z))$$

Penman assumption (a linear Taylor's expansion of saturated tension around the air temperature gives a fourth equation)*

Usually the water budget does not enter in this estimation. One ansatz is to include in C_E also a linear dependence on water stress

$$C_E = \frac{S_g(t)}{S_{max}} C'_E$$

With this in mind:

* Details at http://abouthydrology.blogspot.com/2019/01/material-for-geoframe-winter-school_12.html

The 4 traditional equation in become 5

0 $$\frac{dS_g}{dt} = J^\bullet(t) - Q_g^\bullet(t) - E_T(t) - Q_{sub}(t)$$

1 $$E_T = \rho C_E \frac{\epsilon}{p} \overline{u}(e^*(z_0) - e(z))$$

2 $$R_S = \lambda E_T + H + R_{ll}$$

3 $$H = a\,\rho C c_P \overline{u}\,(T(z_0) - T(z))$$

4 $$e^*(z_0) - e(z) = \delta e(z) + \Delta(T(z_0) - T(z))$$

Schymanski, S. J., & Or, D. (2017). Leaf-scale experiments reveal an important omission in the Penman–Monteith equation. Hydrology and Earth System Sciences, 21(2), 685–706. http://doi.org/10.5194/hess-21-685-2017

One important note is that they should be solved simultaneously. But, if we forget it we can solve the last 4 and then balance the results.

$$2 \quad T_\Delta = \frac{\gamma}{C_E\Delta + aC\gamma}\frac{R_s - R_{ll}}{\rho\overline{u}c_p} - \frac{C_E}{C_E\Delta + aC\gamma}\delta e(z)$$

$$4 \quad e_\Delta = \frac{\gamma\Delta}{C_E\Delta + aC\gamma}\frac{R_s - R_{ll}}{\rho\overline{u}c_p} - \frac{C}{C_E\Delta + aC\gamma}\delta e(z)$$

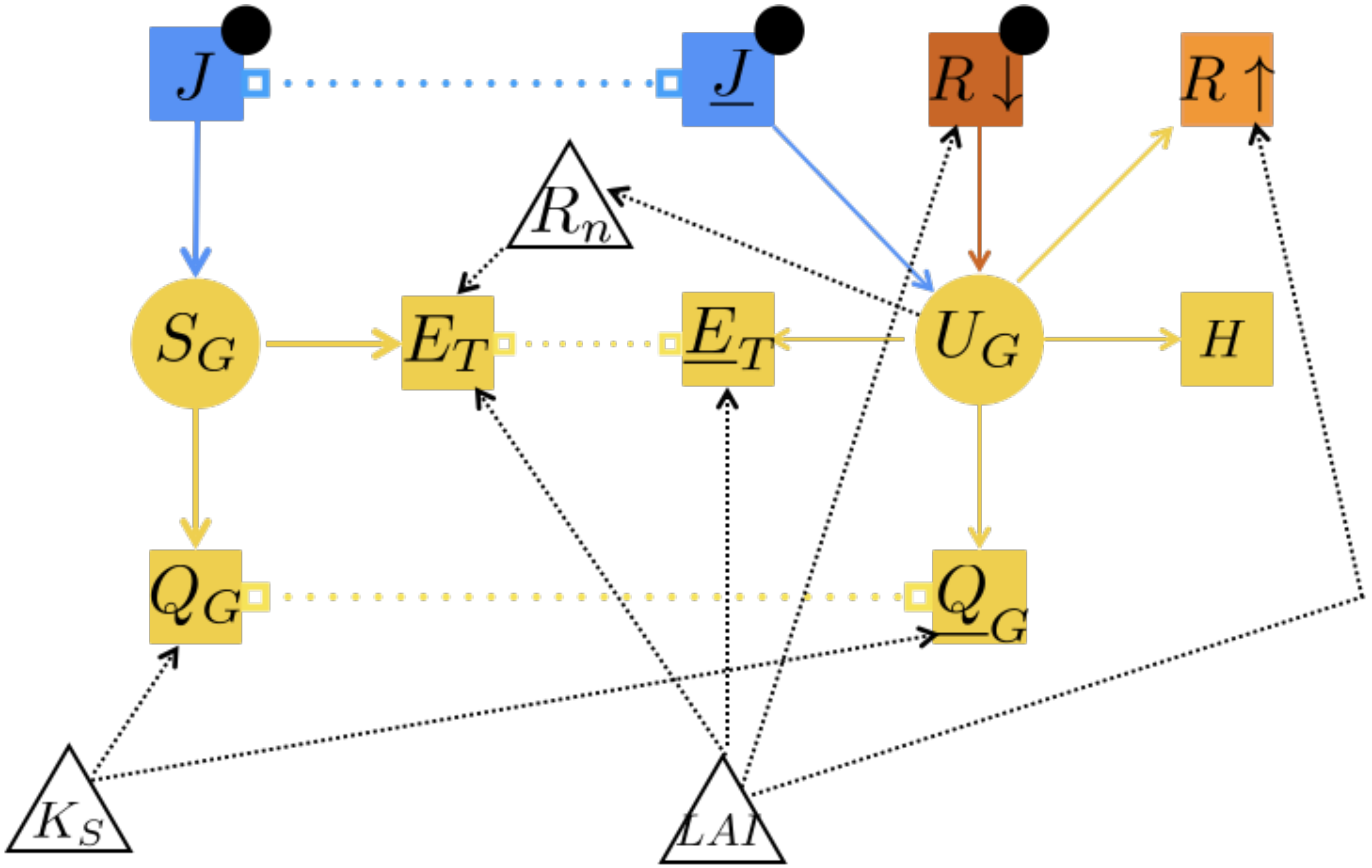
$$3 \quad H = \frac{aC\gamma}{C_E\Delta + aC\gamma}(R_s - R_{ll}) - \rho\overline{u}c_p\frac{aCC_E}{C_E\Delta + aC\gamma}\delta e(z)$$

$$1 \quad \lambda E_T = \frac{C_E\gamma}{C_E\Delta + aC\gamma}(R_s - R_{ll}) - \rho\overline{u}c_p\frac{CC_E}{C_E\Delta + aC\gamma}\delta e(z)$$

$\gamma := \frac{c_p p}{\epsilon\lambda}$ is the so-called psycrometric constant

* Some details at http://abouthydrology.blogspot.com/2019/01/material-for-geoframe-winter-school_12.html

26



A coupled simple model, like the one we solved the equation

Bancheri, M., Serafin, F., & Rigon, R. (2019). The representation of hydrological dynamical systems using the extended Petri Nets (EPN), 1–58.

This solution requires a few new measurements:

- Radiation
- Temperature of air
- horizontal wind velocity
- Vegetation cover

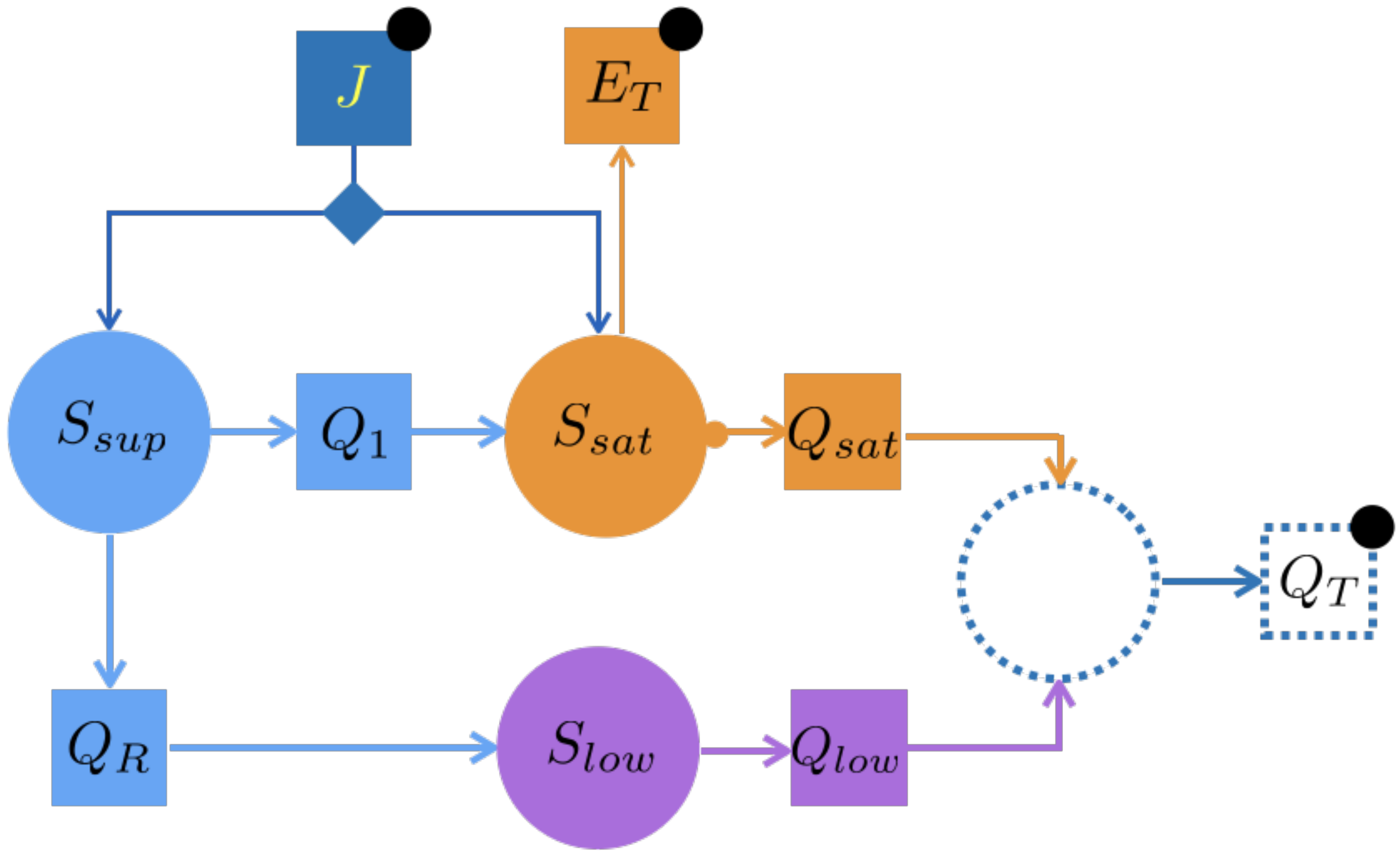
With these we estimate

- Leaf/Soil* temperature
- air vapor pressure
- Evapo-Transpiration
- Turbulent transport of thermal energy

We have to measure more things but we get also more state variables known

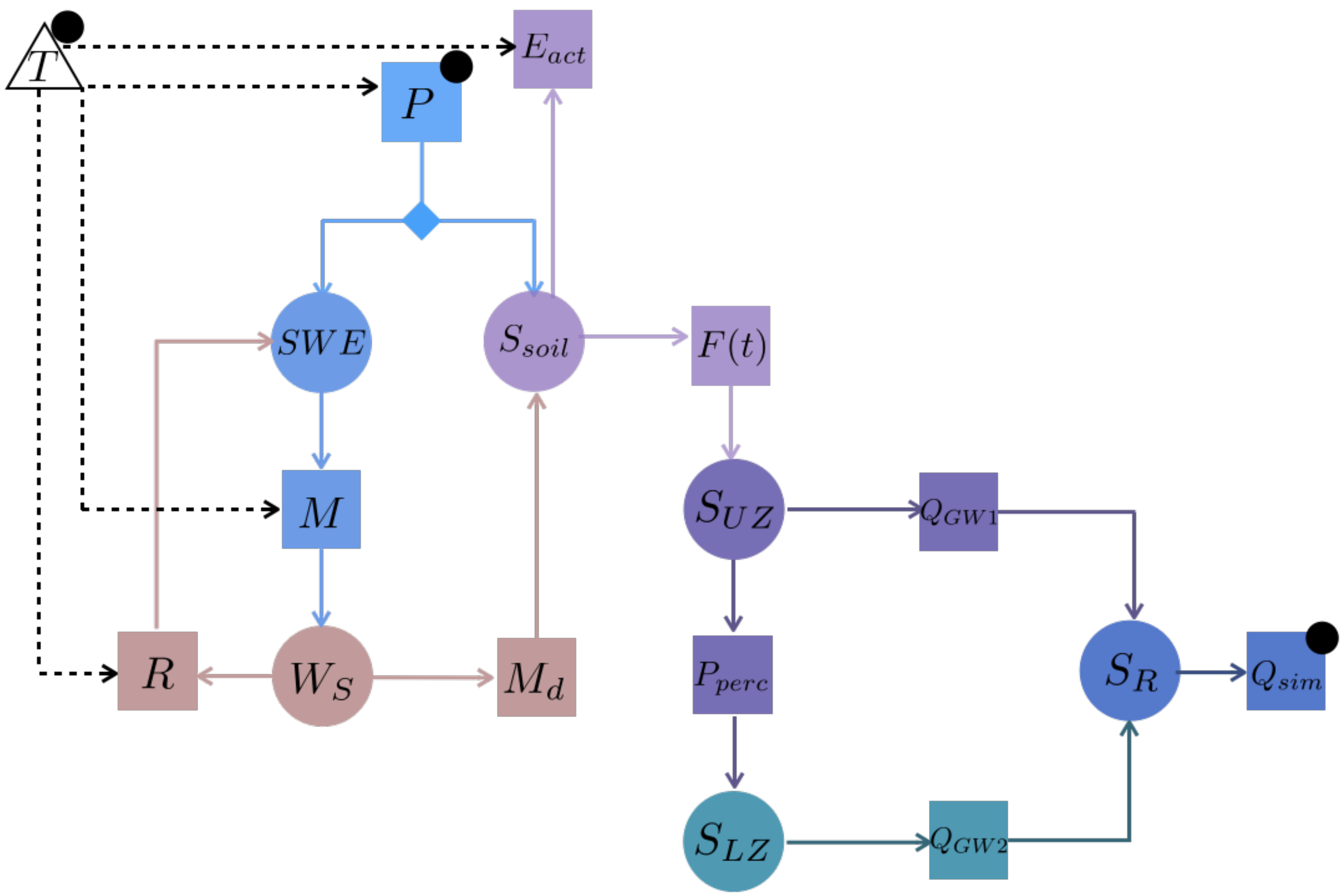
I confess I am hiding quite a few details here

Now, a less than minimal model has more than one reservoirs for HRU



The BST model.

Birkel, C., Soulsby, C., & Tetzlaff, D. (2011). Modelling catchment-scale water storage dynamics: reconciling dynamic storage with tracer-inferred passive storage. Hydrological Processes, 25(25), 3924–3936.

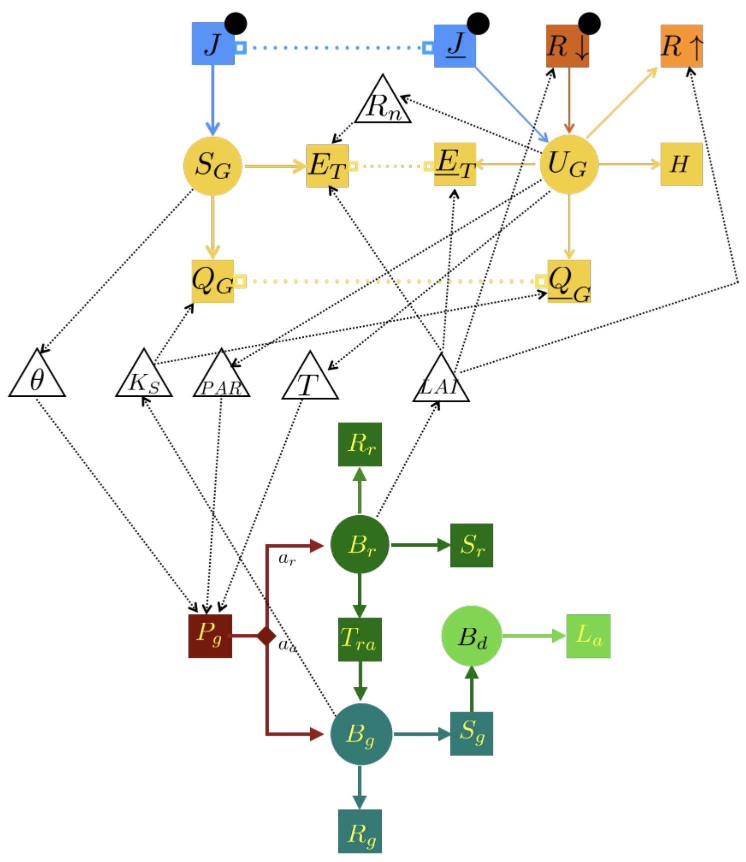


The HBV model.

Seibert, J., & Vis, M. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. Hydrology and Earth System Sciences, 16(9), 3315.

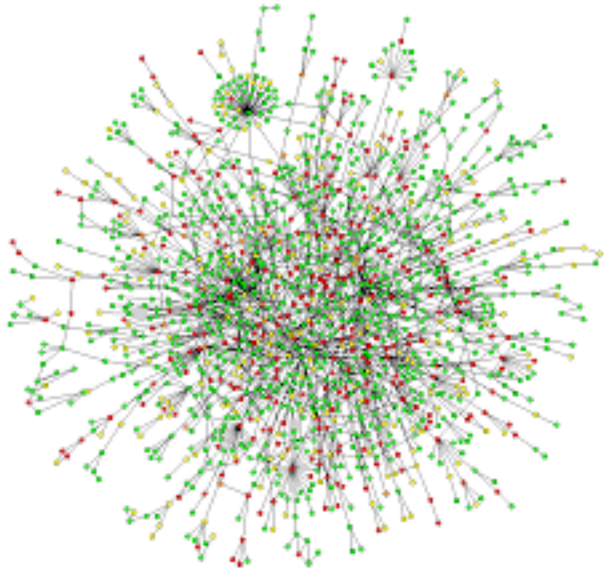
In conclusion, in **contemporary modelling**, all information is certainly absorbed to obtain a **more complete and clear description of the system.** Parameters heterogeneity require it all.

No chance to have too much data !



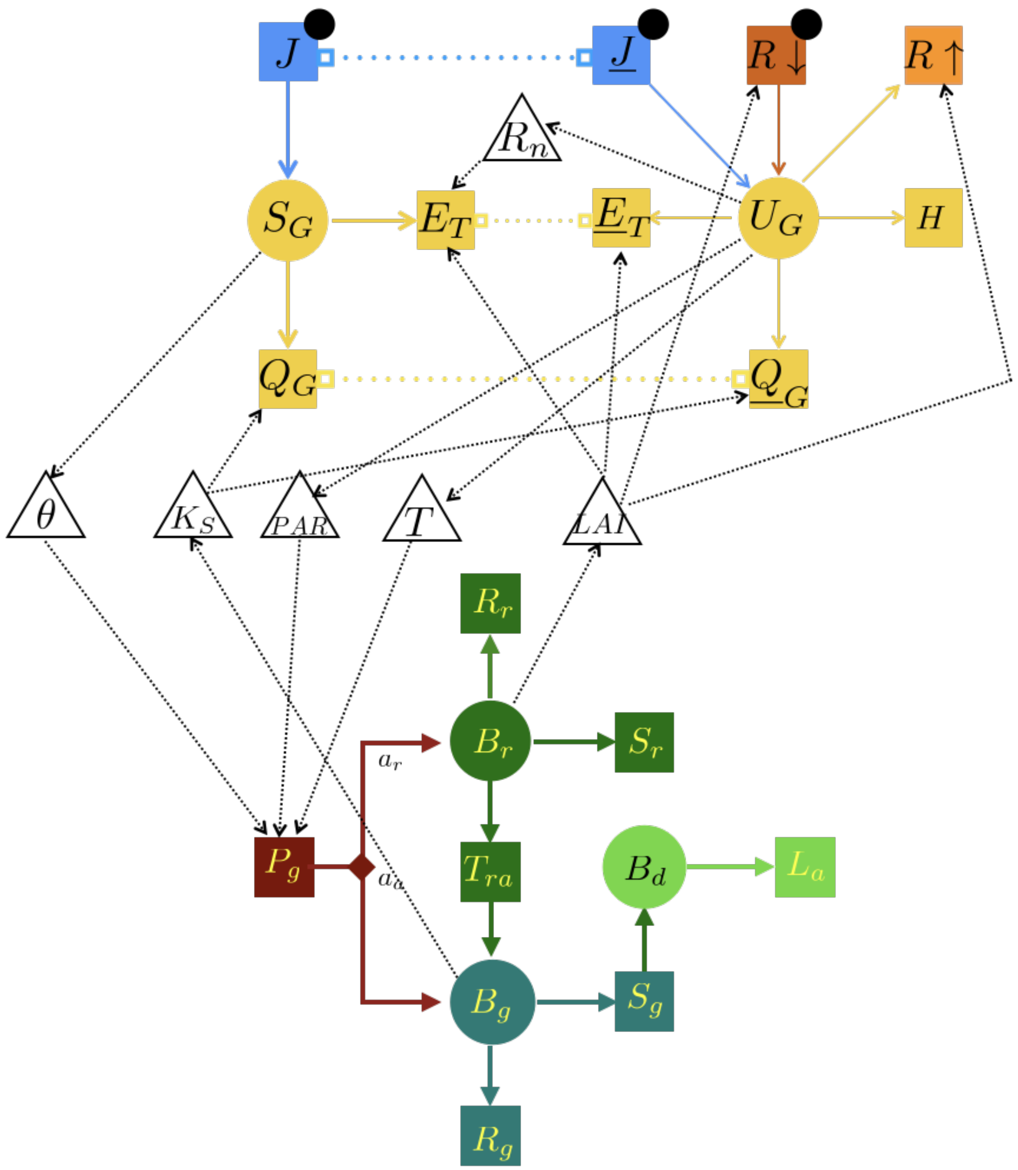
if we want to have a clear view of the system

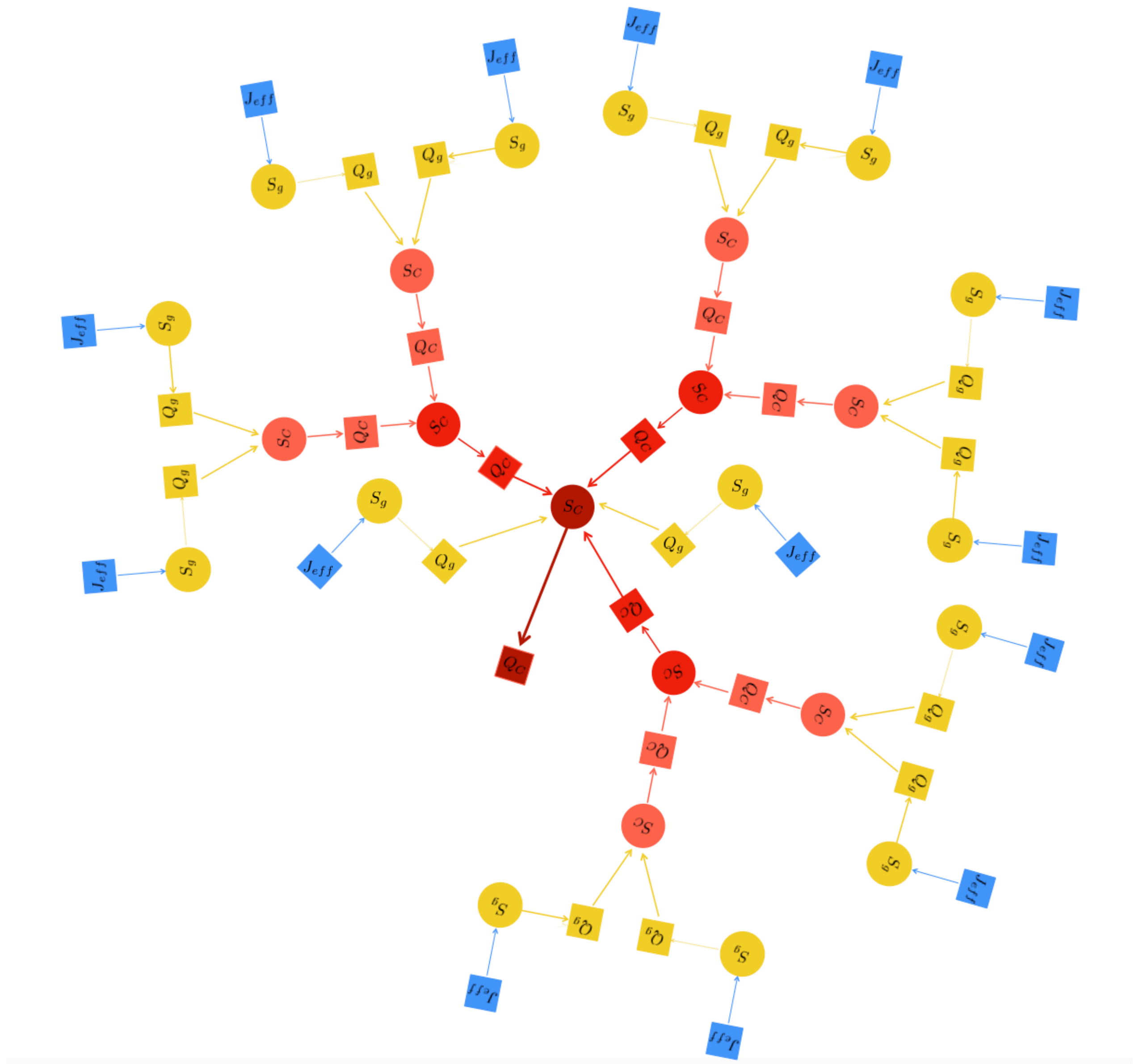
All these models I presented assume some granularity of the processes. In a perfect world this granularity should be obtained by actions that mimics Statistical Mechanics in which laws and fluxes **emerge** from cancellation of degrees of freedom.



In practice we have to rely on appropriately designed Statistics and Calibration, which always depends on Data. Calibration of complicated models is more complicate*.

*Voinov, A., & Shugart, H. H. (2013). "Integronsters," integral and integrated modeling. Environmental Modelling and Software, 39, 149–158. http://doi.org/10.1016/j.envsoft.2012.05.014





Calibration and analysis in portion is possible using the graph-strucured organization of models*

*Serafin, F., (David, O & Rigon, R.) Enabling modeling frameworks with surrogate modelling capabilities and complex networks, Ph.D. Dissertation, 2019

Thursday, January 17, 2019

Pictures from the Winter School on GEOframe

The course for doctoral students, **post docs and young researchers in Hydrology, Forestry,** and related disciplines will cover the simulation of the hydrological cycle of catchments of various sizes with the GEOframe system. They say that all models are wrong but useful. However, with better tools you forecast and decide better.
The course has been completed and, please you can find below all the material, video, papers cited by clicking on the links.

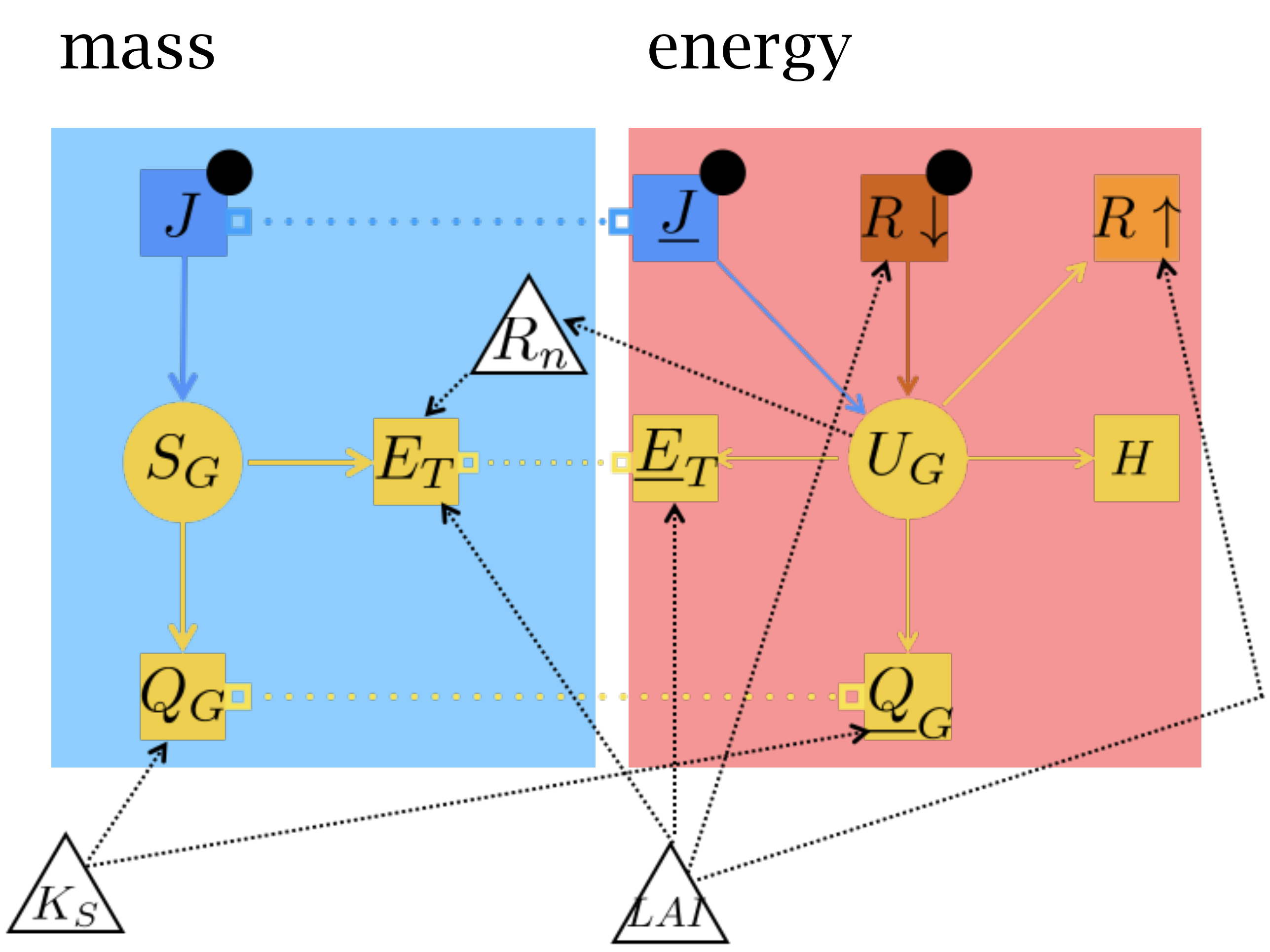
The topics treated has been:

- Tuesday, January 8: Introduction to the course - What is OMS - What is GEOframe - Using Python and Jupyter for visualising Inputs and Outputs
- Wednesday, January 9- Catchments and Hydrologic Response Units delineation
- Thursday and Friday, January, 10 - 11 - Treatment of spatial data and Calibration
- Monday January 14 - Estimation of radiation components
- Tuesday, January 15 - Evaporation and Transpiration
- Wednesday, January 16 - Rainfall-Runoff modelling with various NewAGE components
- Thursday, January 17 - Rainfall-Runoff modelling with various NewAGE components
- Friday, January 18 - Opinions on hydrological modelling. Exercises on assignments.
- Epilogue and thanks

Go to the first day - Go to the last page of the School

http://abouthydrology.blogspot.com/2019/01/pictures-from-winter-school-on-geoframe.html

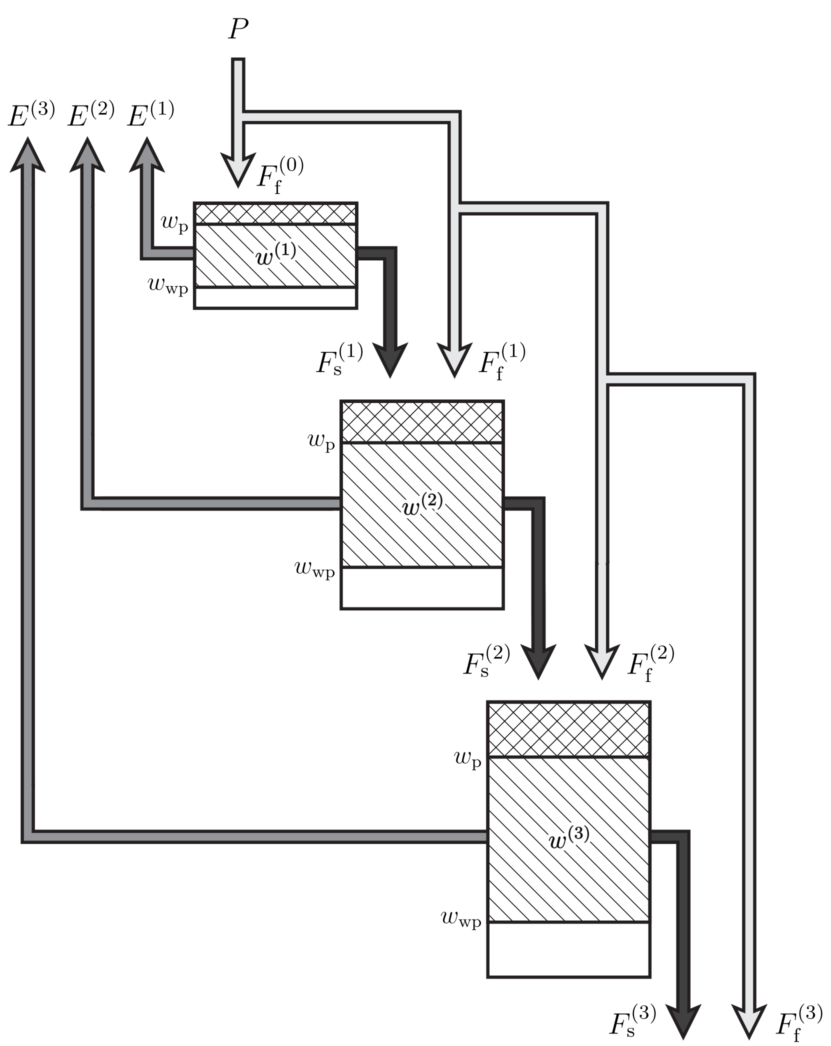
That's why we use conservation laws.

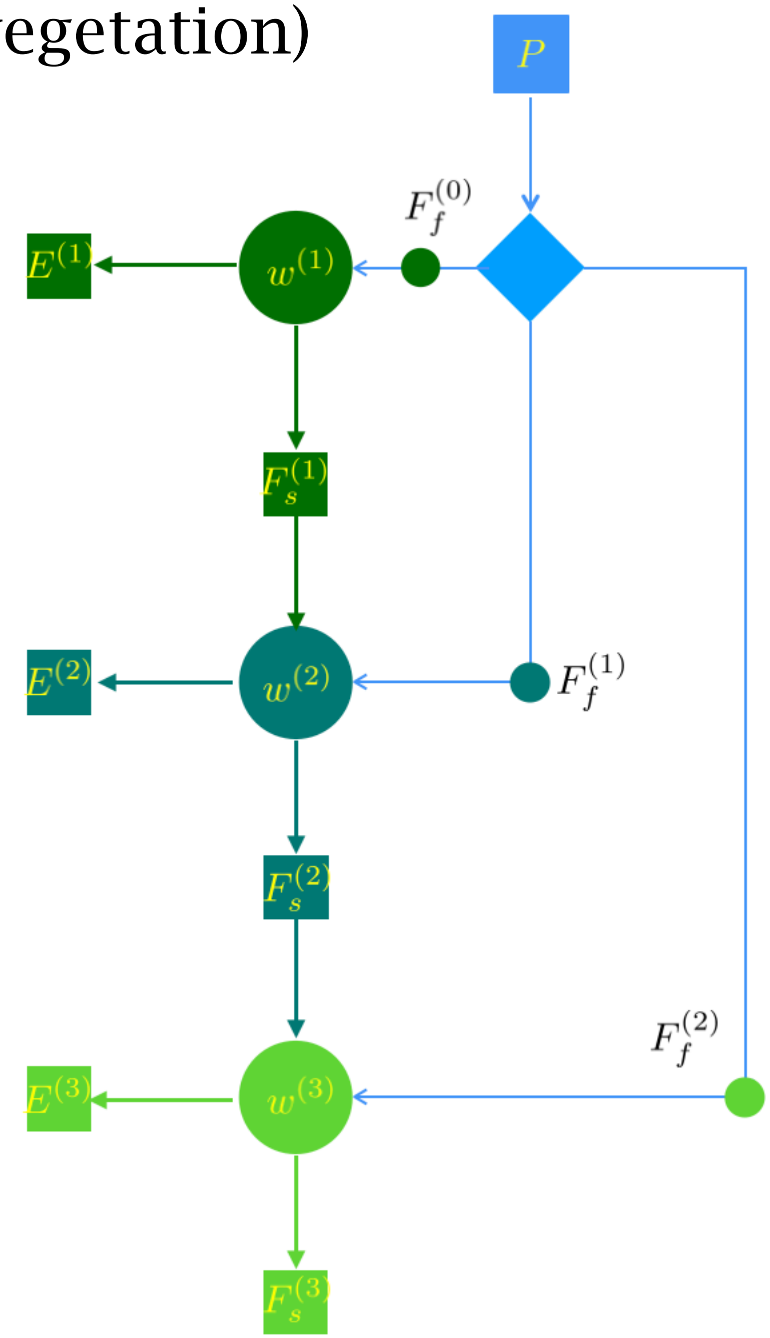


They offer a big constraint on variability

* Oreskes, N., 2003. The role of quantitative models in science. In: Canham, Charles D., Cole, Jonathan J., Lauenroth, William K. (Eds.), Models in Ecosystem Science. Princeton University Press, Princeton, pp. 13e31.

The case of GLEAM v3 (model for tall vegetation)





0-10 cm

10-100 cm

100-250 cm

- lateral fluxes + satellite data of soil moisture (assimilation)

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Férnandez-Prieto, D., et al. (2016). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geoscientific Model Development Discussions, 1–36. http://doi.org/10.5194/gmd-2016-162

- **Models and data** get along.
- **Better data** implies **more refined models**
- More refined models requires more data, **specific data** though.
- With data, **spatial heterogeneity** can be accounted for explicitly.
- **Remote** sensing alone is nothing.
- **Budgets** are what really count, because conservation laws imply constraints on dynamics.

Well, this is not the end, is just the beginning.

- Studies of fluxes of **information and entropy**
- Distinguishing **complication** from **complexity**
- Besides having stuff sorted out **the right way, for the right reasons**,which is the usual bussiness

"While Silberstein (2006) insists that more data are needed and that models without data are not science, Anderson (2008) claims that there is already a deluge of data that can change the way science is done. Google and other similar efforts produce evidence that by analyzing huge arrays of data only, we can actually build new theories based only on correlation, ignoring causation. The new analytical tools that are going to be developed for petabyte computing in the "clouds" will require entirely different approaches to integration than the types of model integration that we were considering so far."*

Anderson, C., 2008. The End of Theory: the Data Deluge Makes the ScientificMethod Obsolete. The Wired Magazine, August. http://www.wired.com/ science/discoveries/magazine/16-07/pb_theory.

Silberstein, R.P., 2006. Hydrological models are so good, do we still need data?Environmental Modeling & Software 21, 1340e1352.

*Voinov, A., & Shugart, H. H. (2013). "Integronsters," integral and integrated modeling. Environmental Modelling and Software, 39, 149–158. http://doi.org/10.1016/j.envsoft.2012.05.014

Find this presentation at

Ulrici, 2000 ?

Other material at

http://abouthydrology.blogspot.com