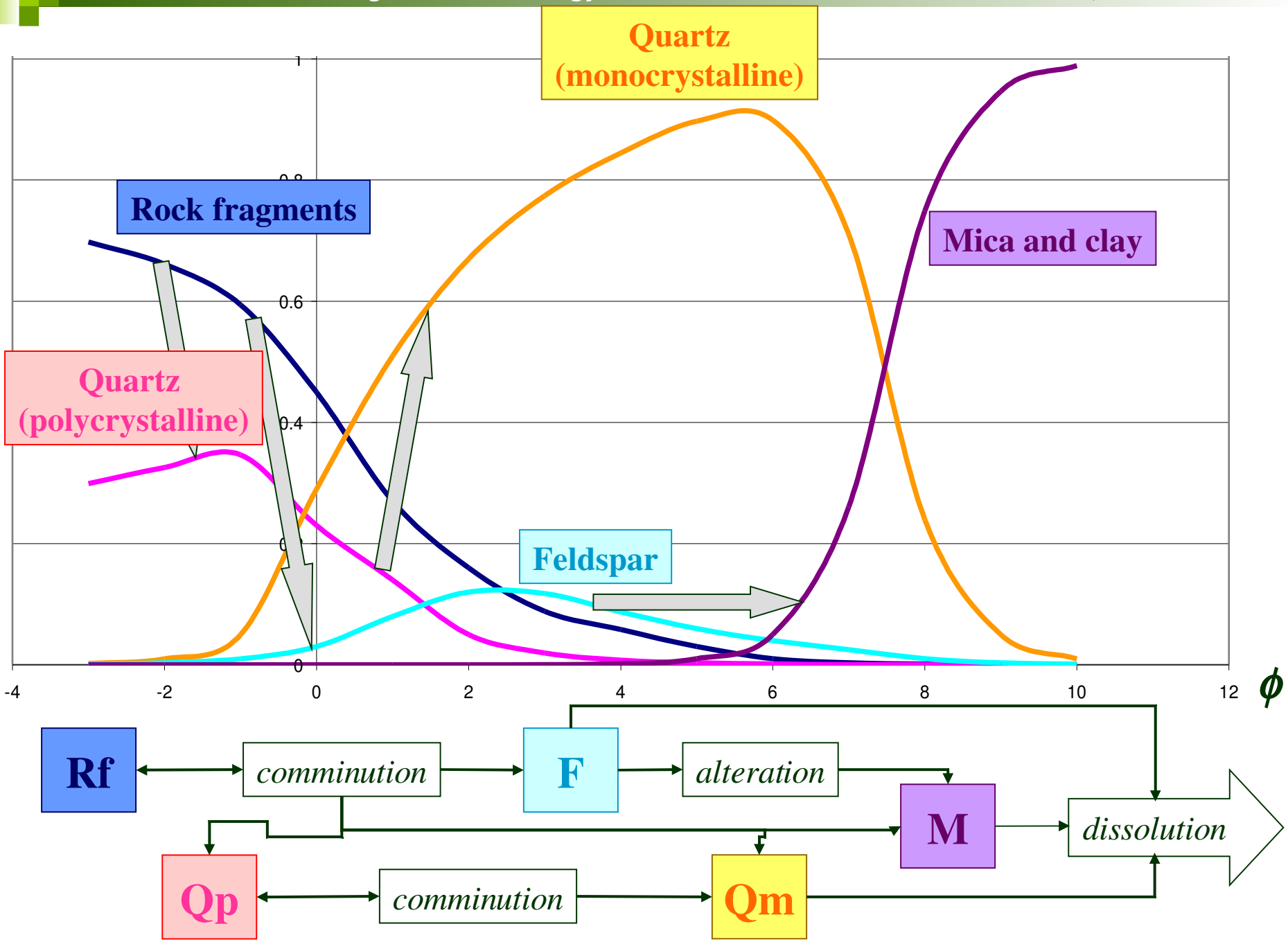
a statistical model of global interaction *grain size-lithology*

regression of a composition with lost values

Seminar zur Sedimentologie und Umweltgeologie– November 2006

raimon.tolosana@geo.uni-goettingen.de


Quartz (monocrystalline)
Rock fragments
Mica and clay
Quartz (polycrystalline)
Feldspar
1
0.6
0.4
0
-4
-2
0
2
4
6
8
10
12
ϕ
Rf
comminution
F
alteration
M
dissolution
Qp
comminution
Qm

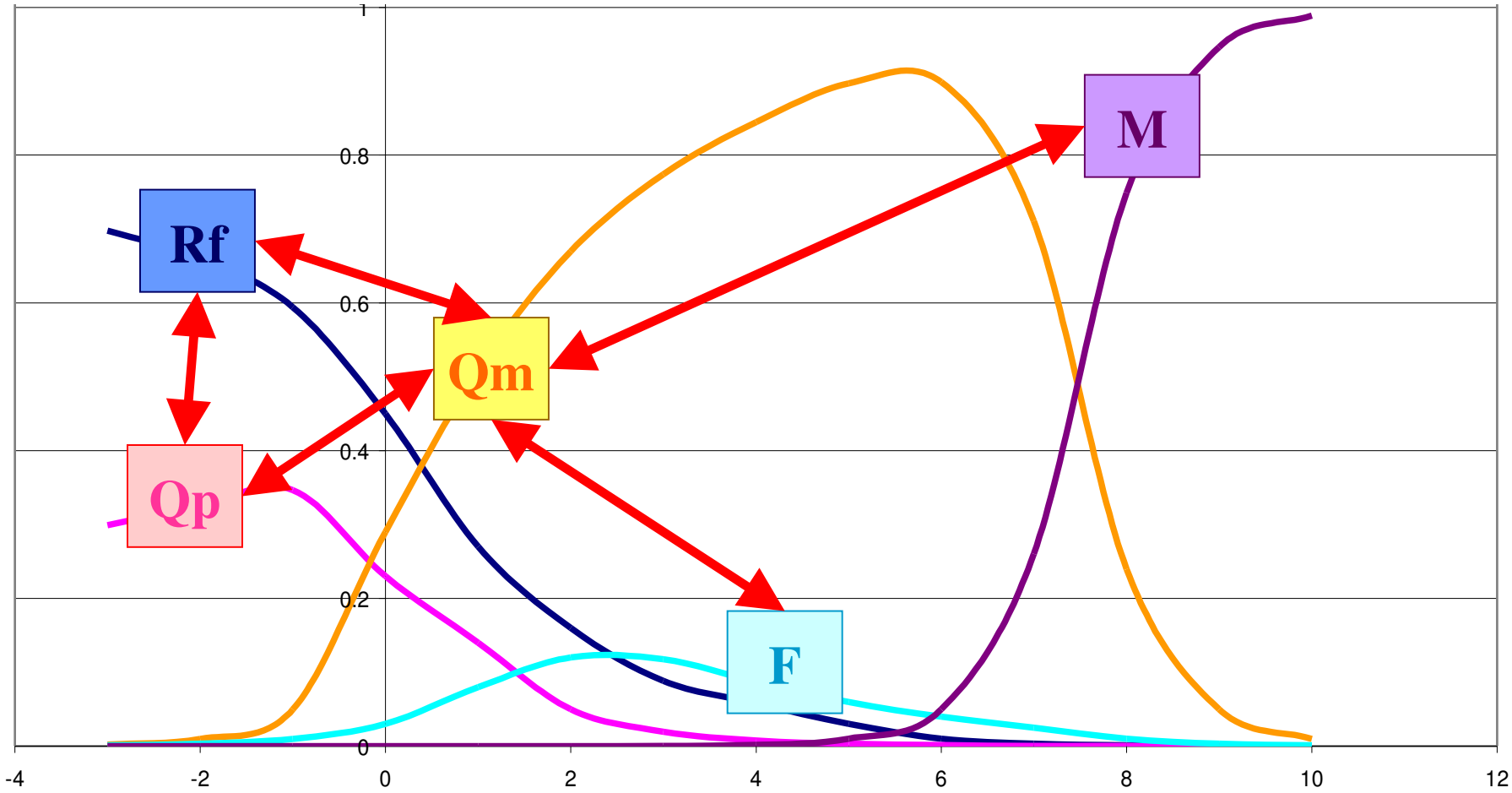
goal

- **x**: composition of D parts, ϕ: grain size $\phi = -\log_2 \frac{d}{mm}$
- explain $\mathbf{x} = f(\phi)$
  - ***BUT***: **x** is a composition (closed), spurious correlation
  - compositions convey ***ONLY*** relative information (ratios)
  - all relative information is contained in any exhaustive set of D–1 log-ratios (Aitchison, 1986)
  4. regress each lr(**x**) on ϕ
  5. invert coefficients, predictions, etc… to compositions
- lots of values below detection limit (***BDL***)!!!
- strategies to follow?

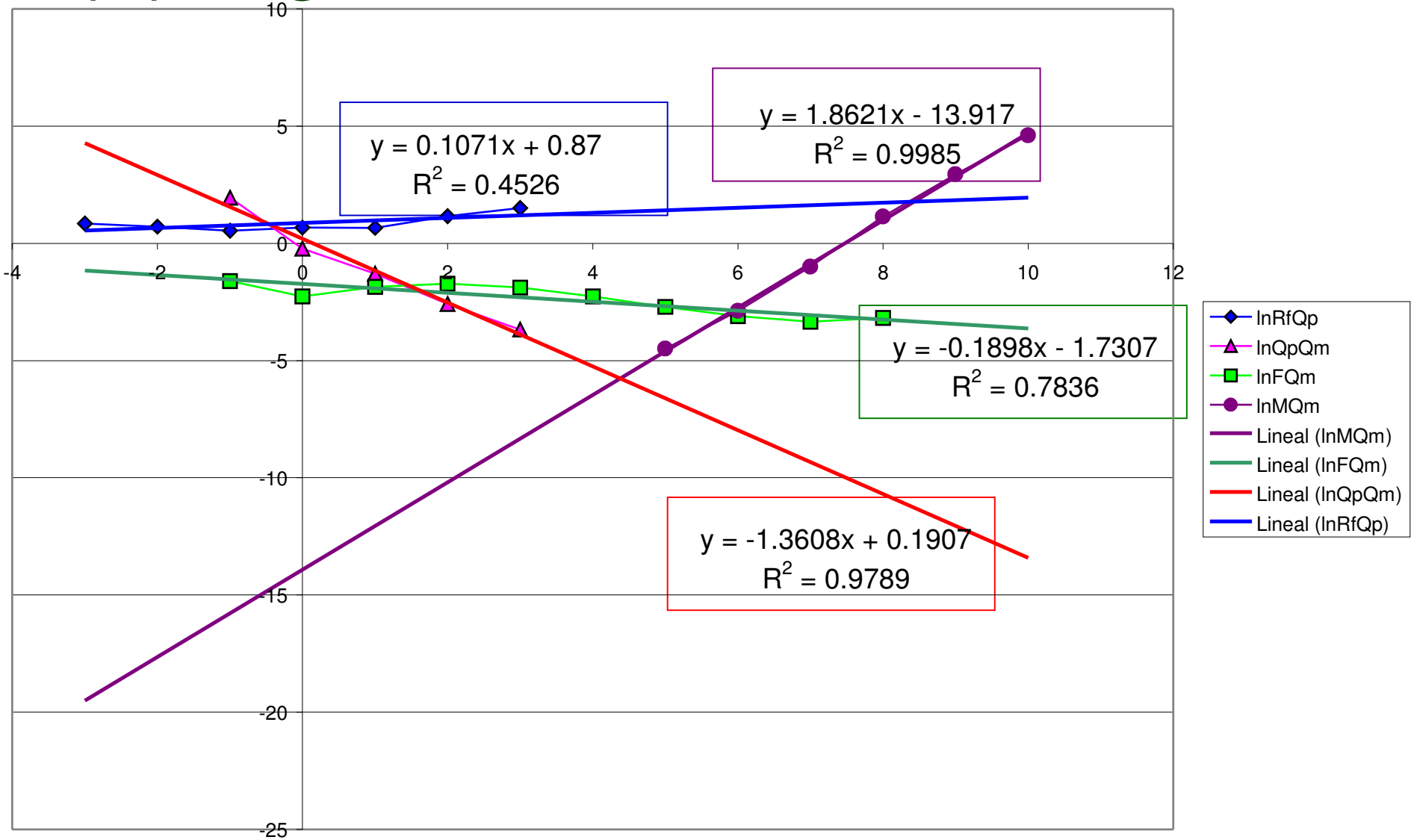
possible methods

1. choosing a "*basis*" (exhaustive set of log-ratios)
   - so that most BDLs *go* together:
     - ratio of two BDLs is a missing value (MV)
     - MVs may be (more) safely ignored
     - ignoring BDLs induces bias (overestimation)
2. replacement strategies (Martín-Fernàndez, 2000)
   - by 2/3 of the DL → by an exponential tail

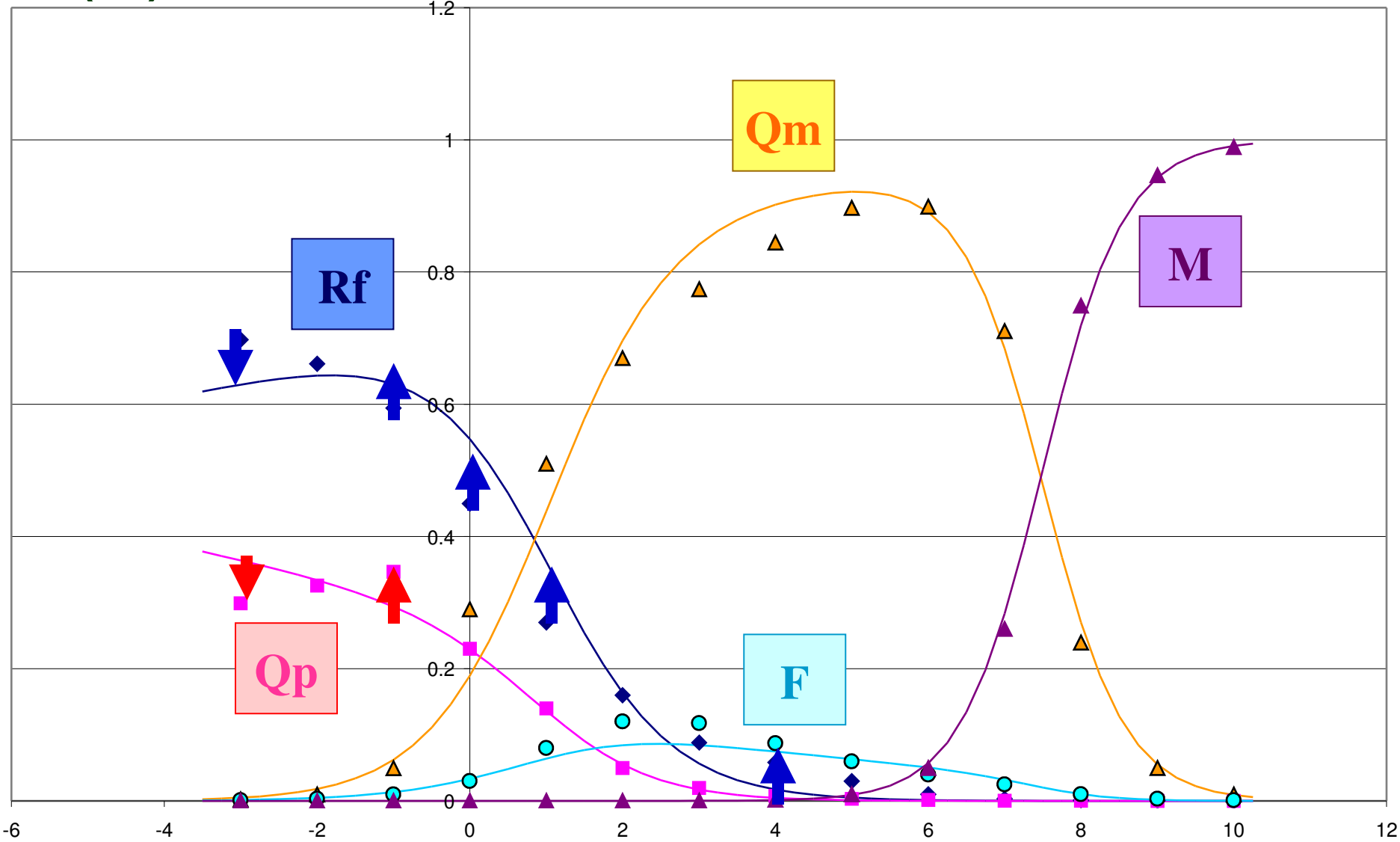
(1) choosing a *basis*



(1) regression on lr basis



(1) results with lr basis



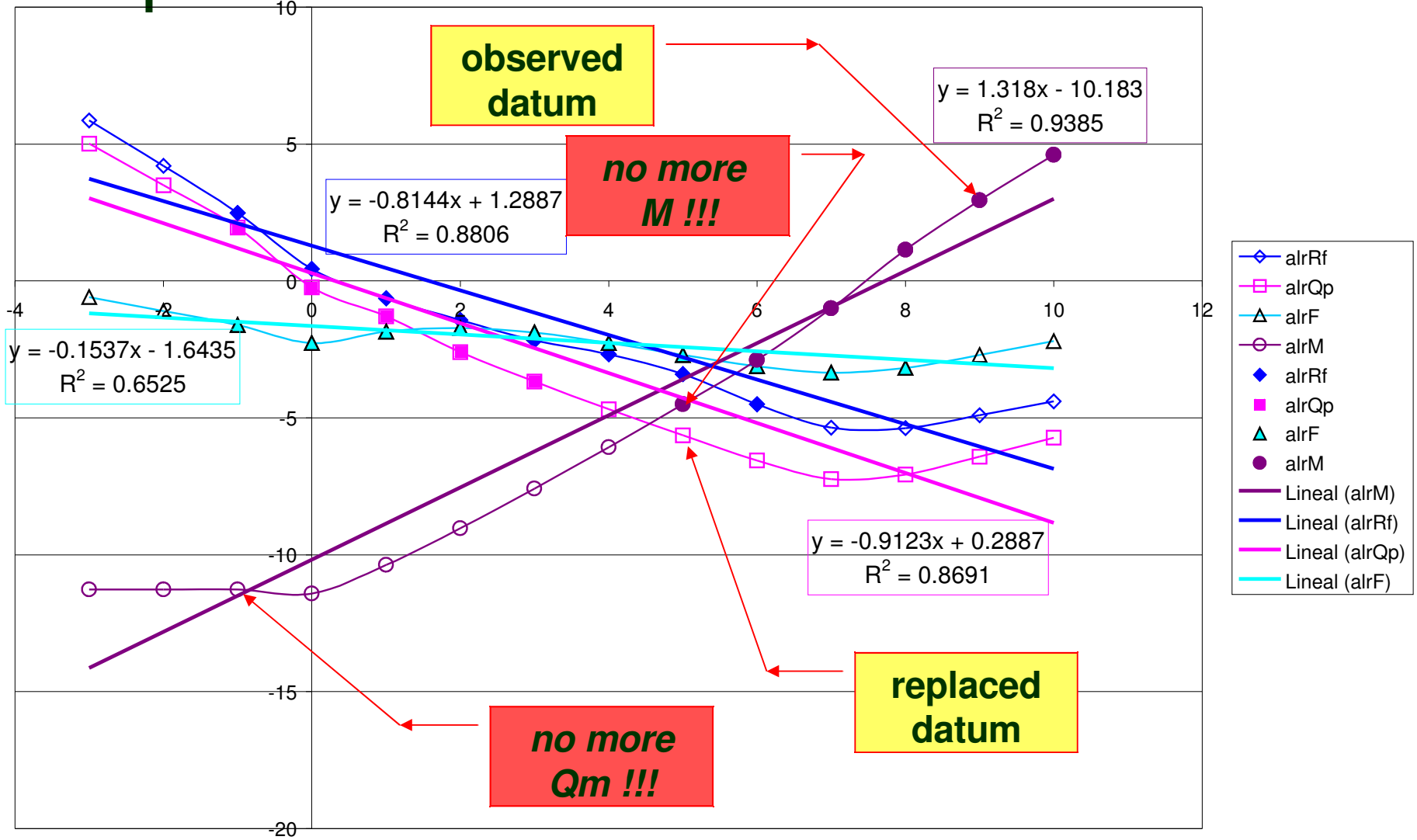
(2) *replacement*

- *original idea*: put 2/3 of detection limit

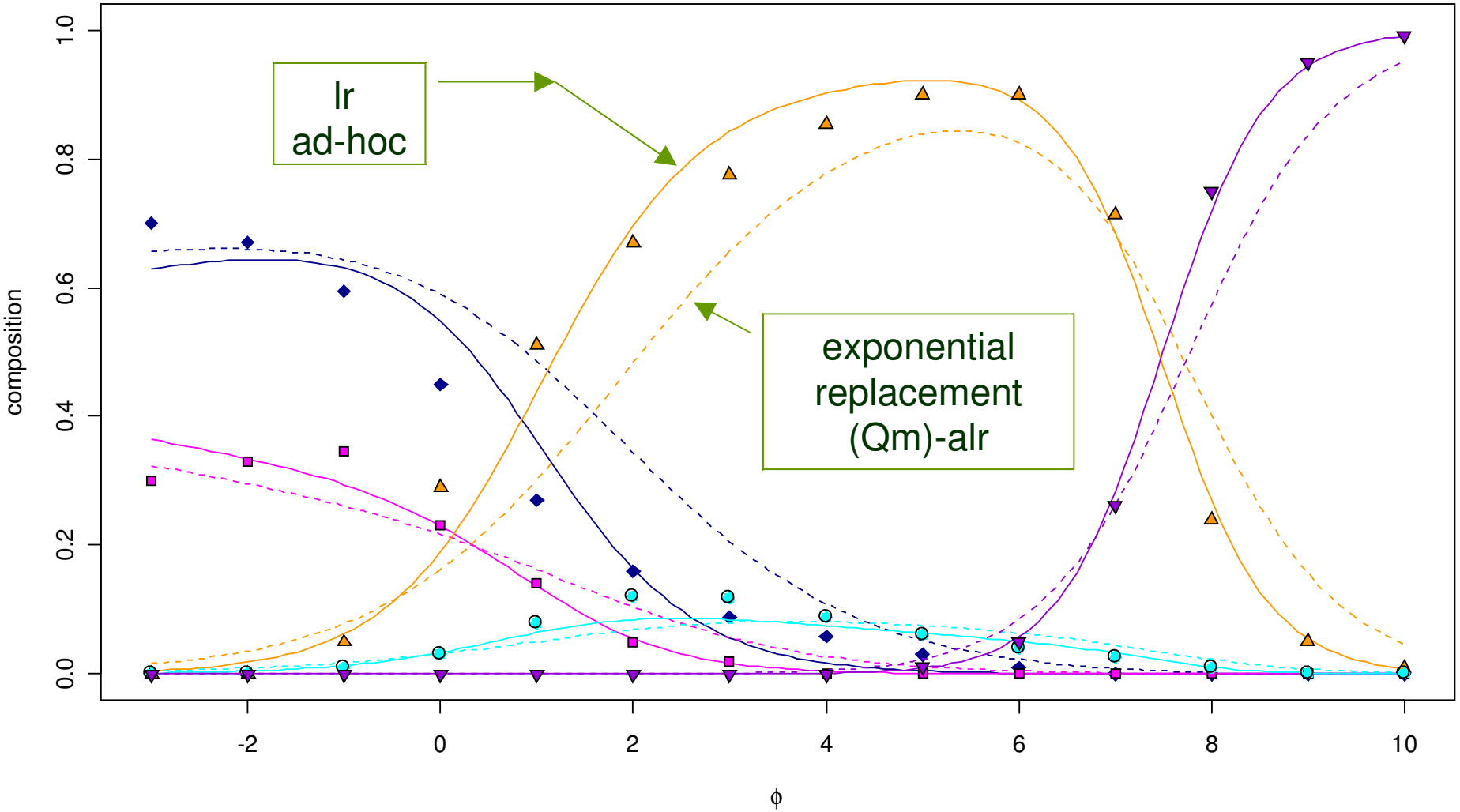
b) *reworked idea*:
   - put an exponential tail
   - following the trend of last 2-3 observations
   - so that it is always below 1%

- regress an arbitrary set of *lr* (*alr*)

(2) regression on exponentially replaced data



comparison 1-2

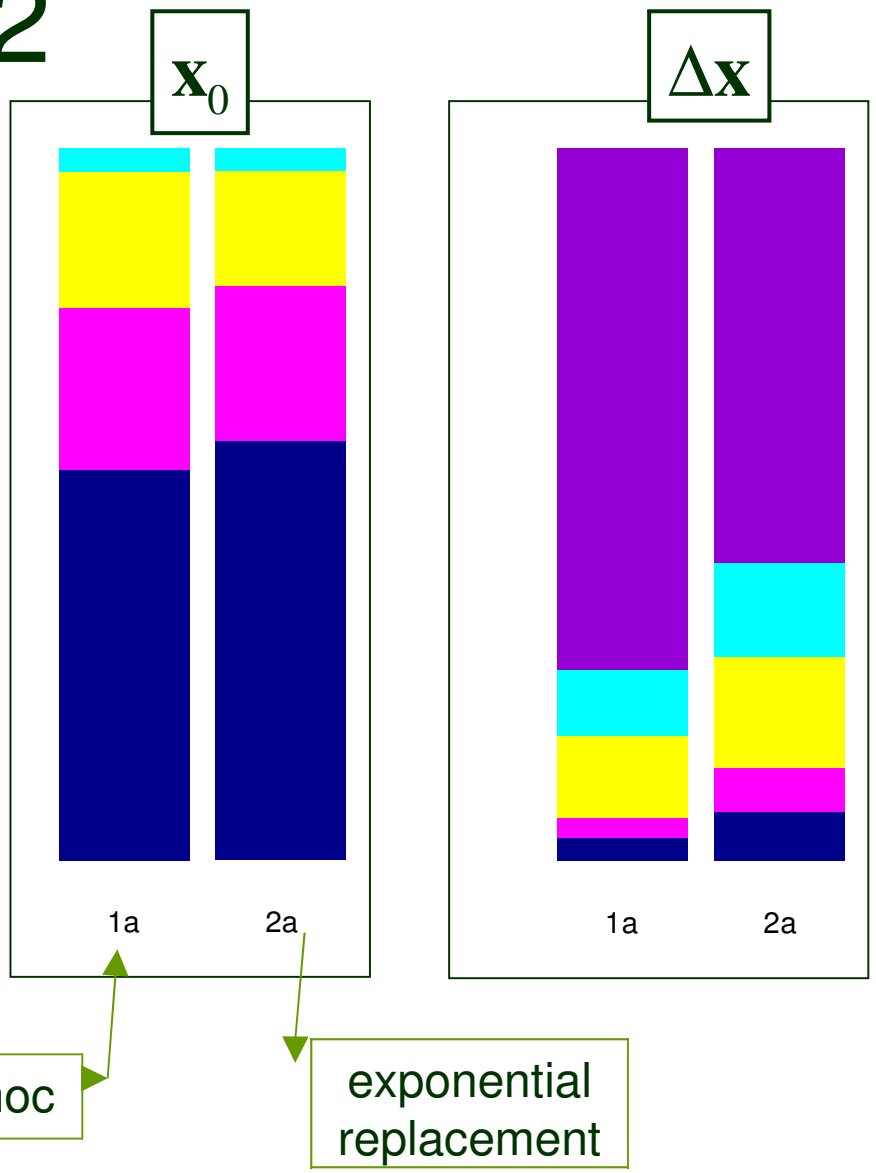


comparison 1-2

- pooling all regression coefficients together:

  $\mathbf{x} = \boldsymbol{C}(\mathbf{x}_0 \cdot \Delta\mathbf{x}^{\phi})$

- $\mathbf{x}_0$: *initial* composition (for ϕ=0, coarse sand)
- $\Delta\mathbf{x}$: compositional variation when ϕ increases a unit (*"splitting" a grain in 8 pieces*)



conclusion

- a complex superposition of processes could be modeled with a one-dimensional Aitchison process (or *balance of processes*):
  - *original composition*: 70% Rf + 30%Qp ($\phi = -3$)
  - *slope*: M >>>> Qm > F > (Rf ≥ Qp)
- ***NOT*** possible with raw data
  - ***solution***: work with an exhaustive set of log-ratios,
    - standard (alr, clr, ilr) or build your optimal one, which can (*should*) be interpretable
    - attends to closed character of data (***NOT*** the system)
- replacement strategies are dangerous/misleading