1st Megha-Tropiques G.V. Workshop on
Rainfall Products Validation and Applications

26 September 2013

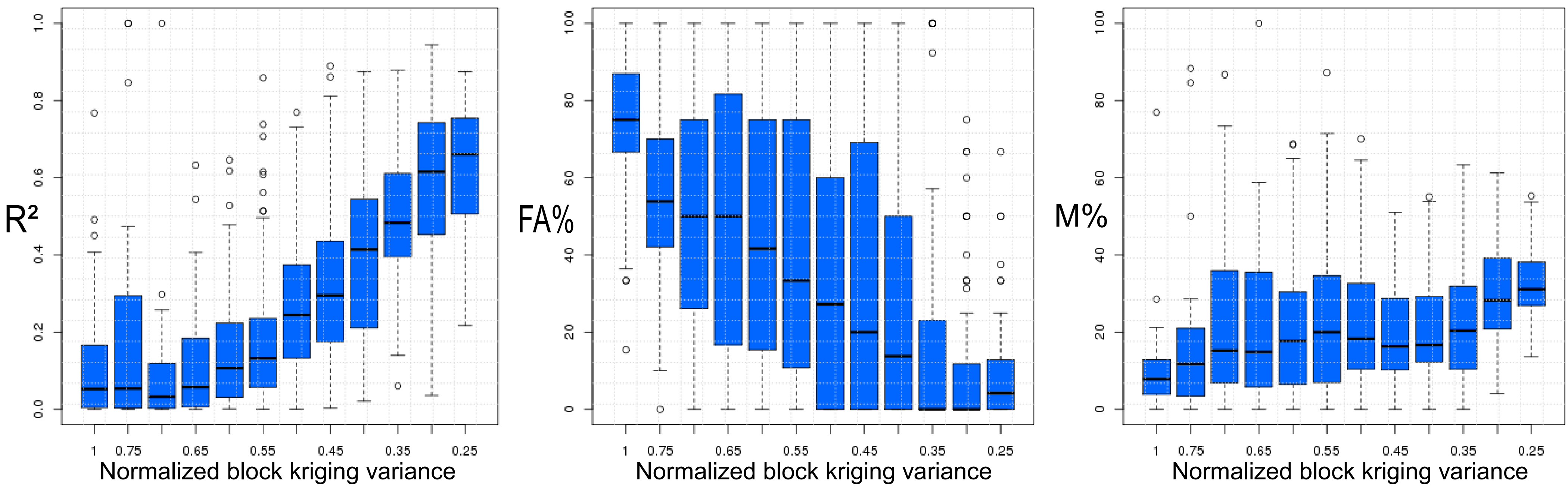
How to account for rain gauges and kriging uncertainties when assessing satellite products

Clément Guilloteau[1], Marielle Gosset[1], Remy Roca[2], Philippe Chambon[3], Guillaume Quantin[4], Theo Vischel[4]

[1] GET, Observatoire Midi Pyrénées, Toulouse
[2] LEGOS, Observatoire Midi Pyrénées, Toulouse
[3] Meteo-France, Toulouse
[4] LTHE, Grenoble

How does network density affects validation scores

TMPA 3B42v7 compared with rain gauges
South America, Oct. Nov. Dec. 2011, 1° - 1day



Denser network => Better correlation / Lower false alarm rate
=> Higher miss rate (due to better detection of low rain rates)

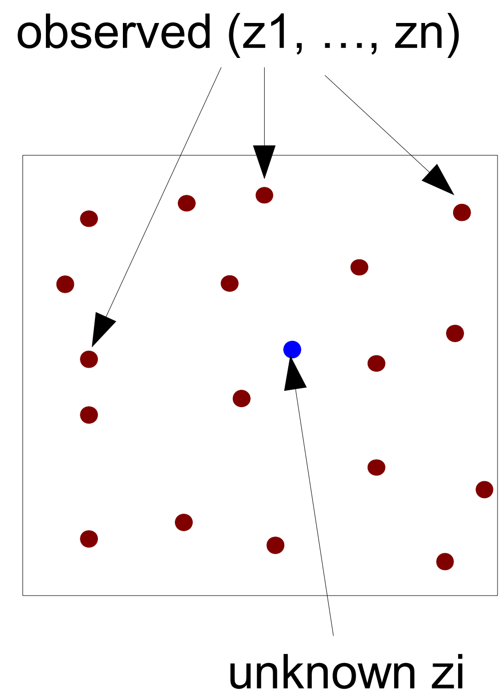
Ordinary kriging, which error model ?

- Rain Z(x,y,t) is a random process with E[Z]=m

- Ordinary kriging theory :

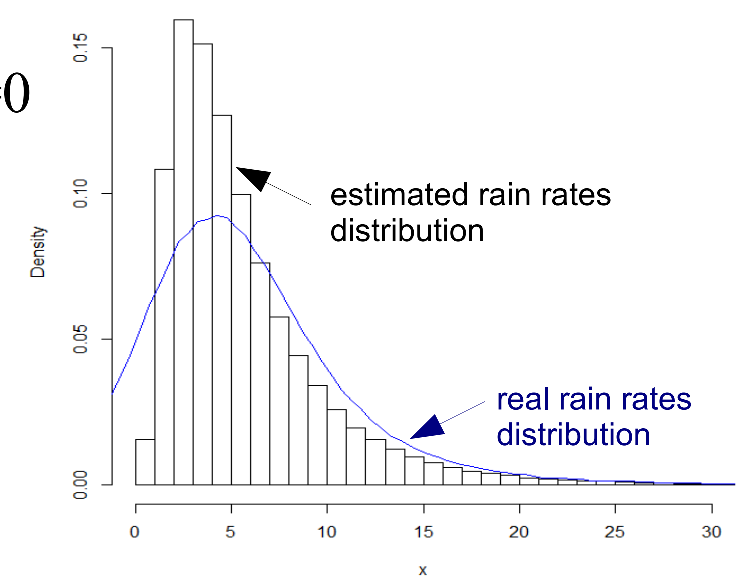
$\hat{z}_i = E[\, z_i \mid z_1, \dots , z_n]$

$E[\hat{Z}] = m = E[Z]$ => Unbiased (true if the set of observed values (z1, …,zn) is large enough and representative)



=> Berkson error model : $z = \hat{z} + \varepsilon$ with $E[\varepsilon]=0$

$var[\hat{Z}] < var[Z]$ "smoothing" effect



Areal mean estimation by local block kriging

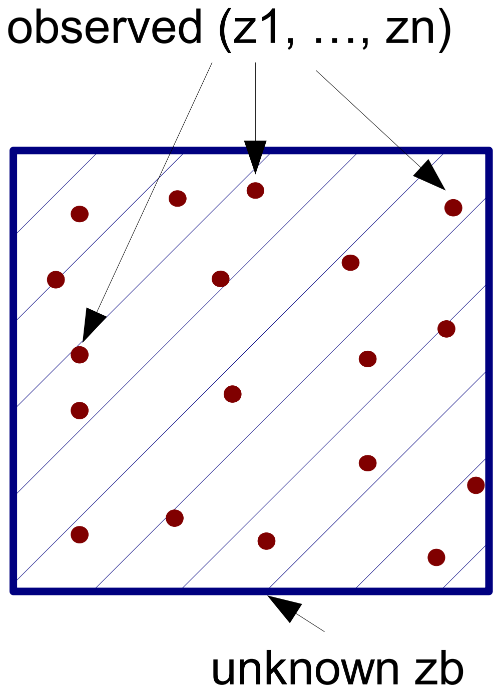
- If all observed values are inside the block, $z_b = m$

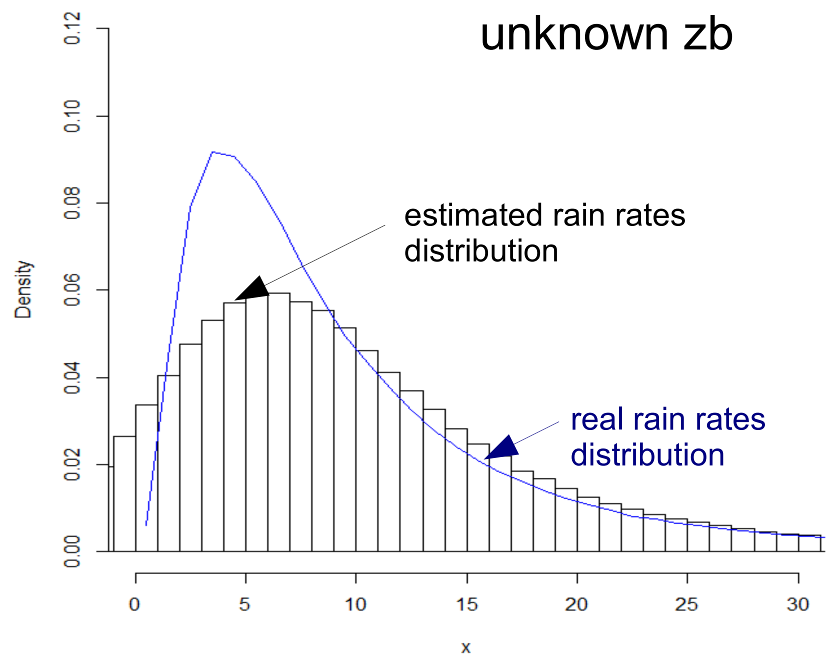
- Equivalent to a sample mean estimation

  => Classical error model : $\hat{z}_b = z_b + \varepsilon$ with $E[\varepsilon]=0$

  $var[\hat{Z}_b] > var[Z_b]$ "noising" effect

- Epistemic uncertainty (such as variogram unaccuracies) is also prone to generate classical additive errors.

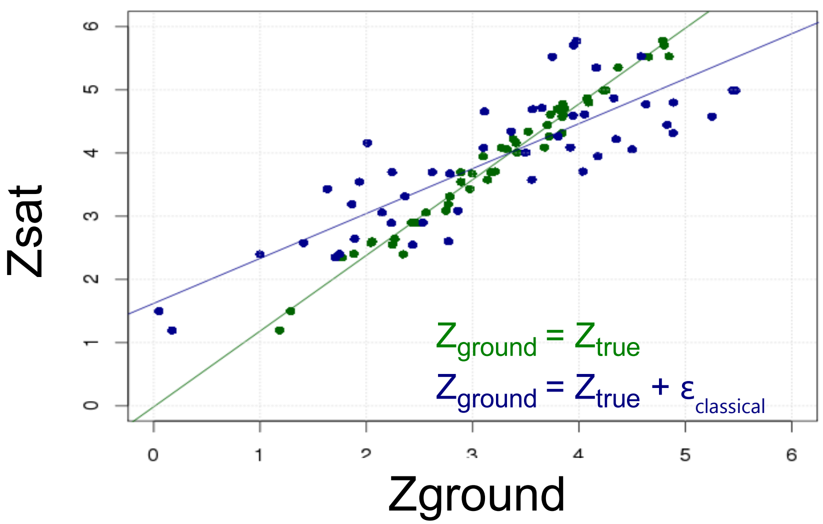




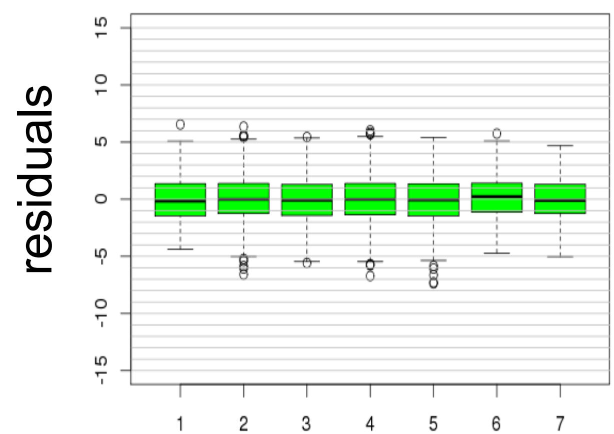
Linear regression between satellite and ground

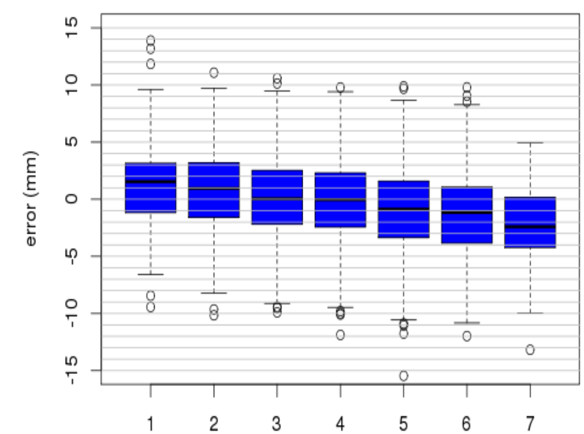
- Effect of classical additive error:

Linear regression Zsat vs Zground

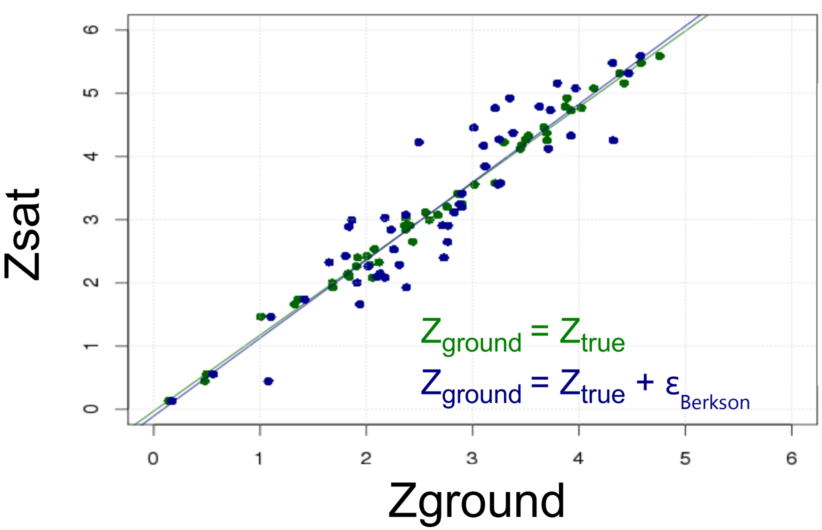


Residuals ($Z_{sat} - Z_{ground}$) distribution by class of Z_{ground}

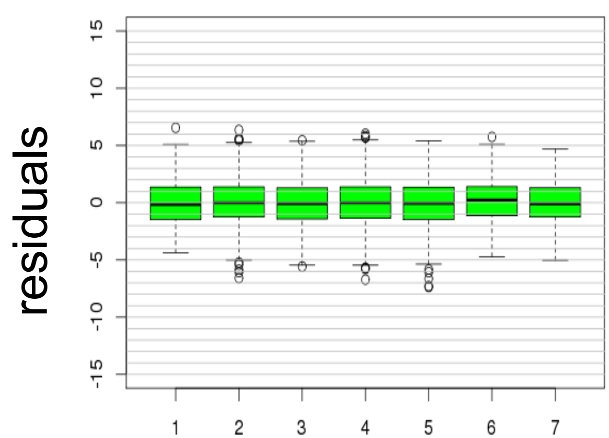




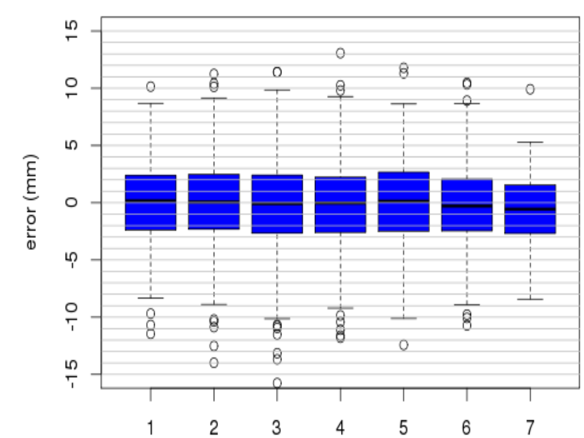
- Effect of Berkson error:



$Z_{ground} = Z_{true}$



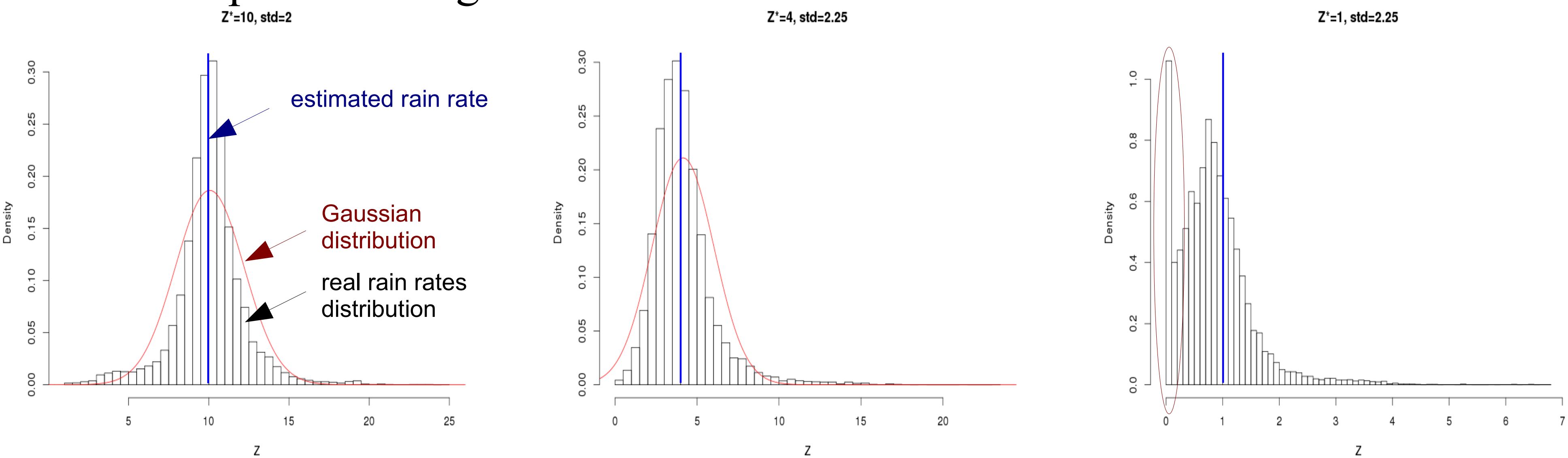
$Z_{ground} = Z_{ground} + \varepsilon$



What is ε PDF ?

- When ordinary kriging hypothesis are verified (second order stationarity), var(ε) is known.
- When Z is normally distributed $f(\varepsilon_i \mid z_1, \ldots, z_n)$ is a gaussian.
- When Z is non-gaussian ?

cross-comparison : daily 1 km² rain rates, (from X-port radar, Niger) sub-sampled and kriged



Small relative error:
non-gaussian
Kurtosis > 3 skew ≈ 0

Large relative error:
non-gaussian
Kurtosis > 3 skew > 0

Small rain rates:
Intermittency effect

nb : Gaussian anamorphosis on rain rates enables to retrieve the simple gaussian case.

Linear regression with errors, a bayesian method

- Kelly/Roca linear regression model :

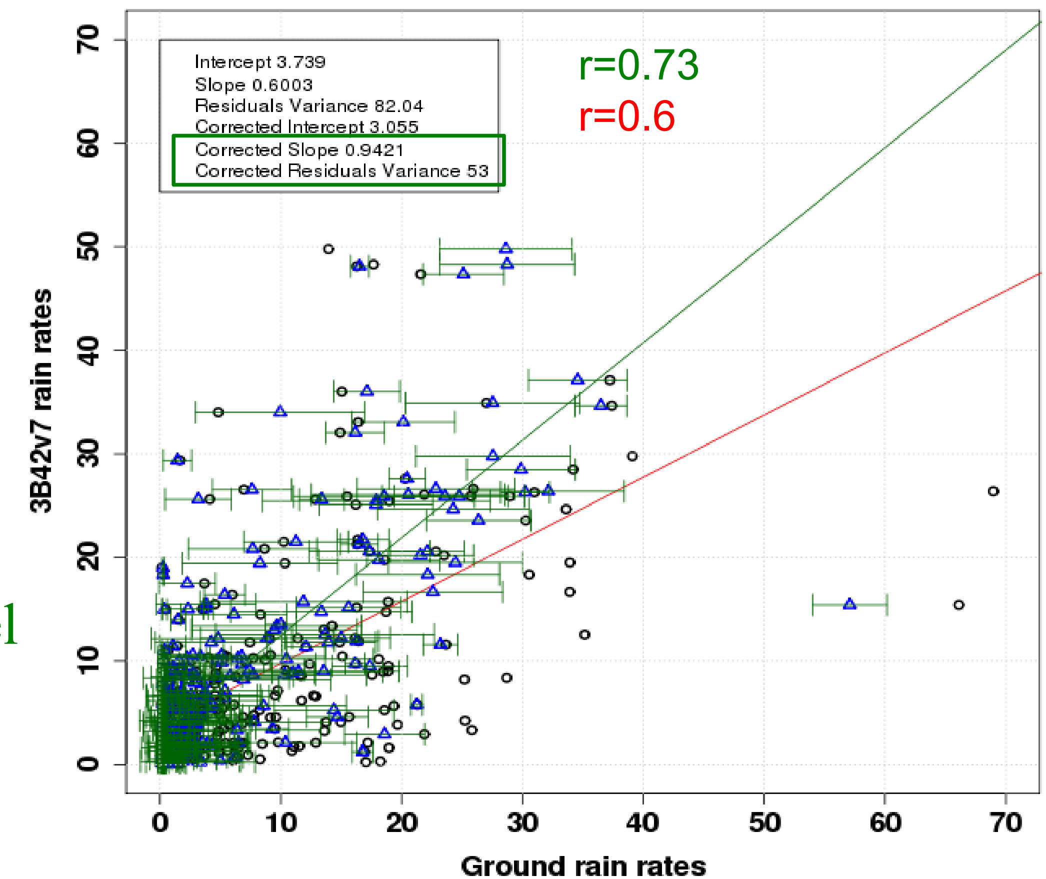
$\hat{Z}_{ground} = Z + \epsilon_{ground}$ (classical error) $\hat{Z}_{sat} = \alpha+\beta*Z + \epsilon_{sat}$

- α, and β estimated with minimum σ_{sat} constraint

$Z = \hat{Z}_{kelly} + \epsilon_{kelly}$ (Berkson error)

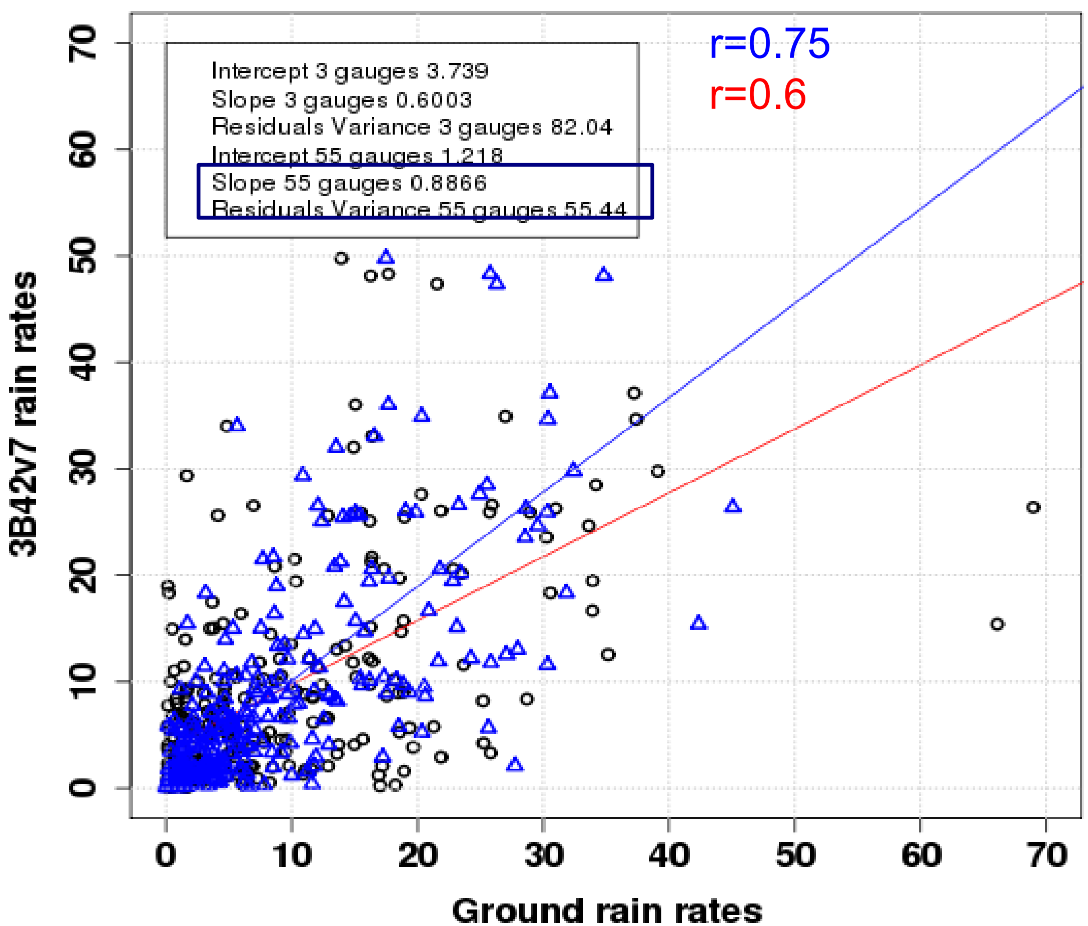
Exemple : TMPA 3B42v7 against 3 gauges network, 1°-1day Niger, 2006-2010

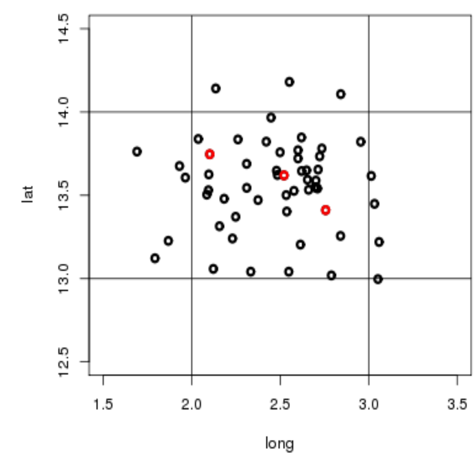
-Red: classical linear regression
-Green: regression with error model



Linear regression with errors, a bayesian method

- Validation : TMPA 3B42v7 against AMMA-CATCH Niger super-site (55 gauges).





Ground errors for the full network are considered negligible.

-Red: linear regression 3 gauges
-Blue: linear regression 55 gauges

Conclusion

- Kriging error Є distribtution is not obvious when Z is not gaussian.

- The pdf of Є is far more informative than it's variance only:
  - to compute rain/no rain probability
  - to compute exeedance probability
  - for bayesian analysis

Ref :

Kelly B. C., Some aspects of measurement error in linear regression of astronomical data, Astrophysical Journal, 2007

Carroll R., Ruppert D., Stefanski L., Crainiceanu C., Measurement error in nonlinear models, Champman & Hall/CRC

A resolution of the problem : gaussian anamorphosis

- Apply a transformation on observed rain rates to obtain a gaussian distribution : $W = \Phi(Z)$
- The PDF of the real rain rates $f(z_i \mid \hat{w}_i, \sigma_i)$ can be retrieved by a change of variable from $f(w_i \mid \hat{w}_i, \sigma_i)$ or by Monte Carlo simulations.

A simple anamorphosis: $W = \sqrt{Z}$

- $\sqrt{Z}$ distribution may be "lesss non-gaussian" than Z distribution

=> Kriging error on $\sqrt{Z}$ is reasonably gaussian