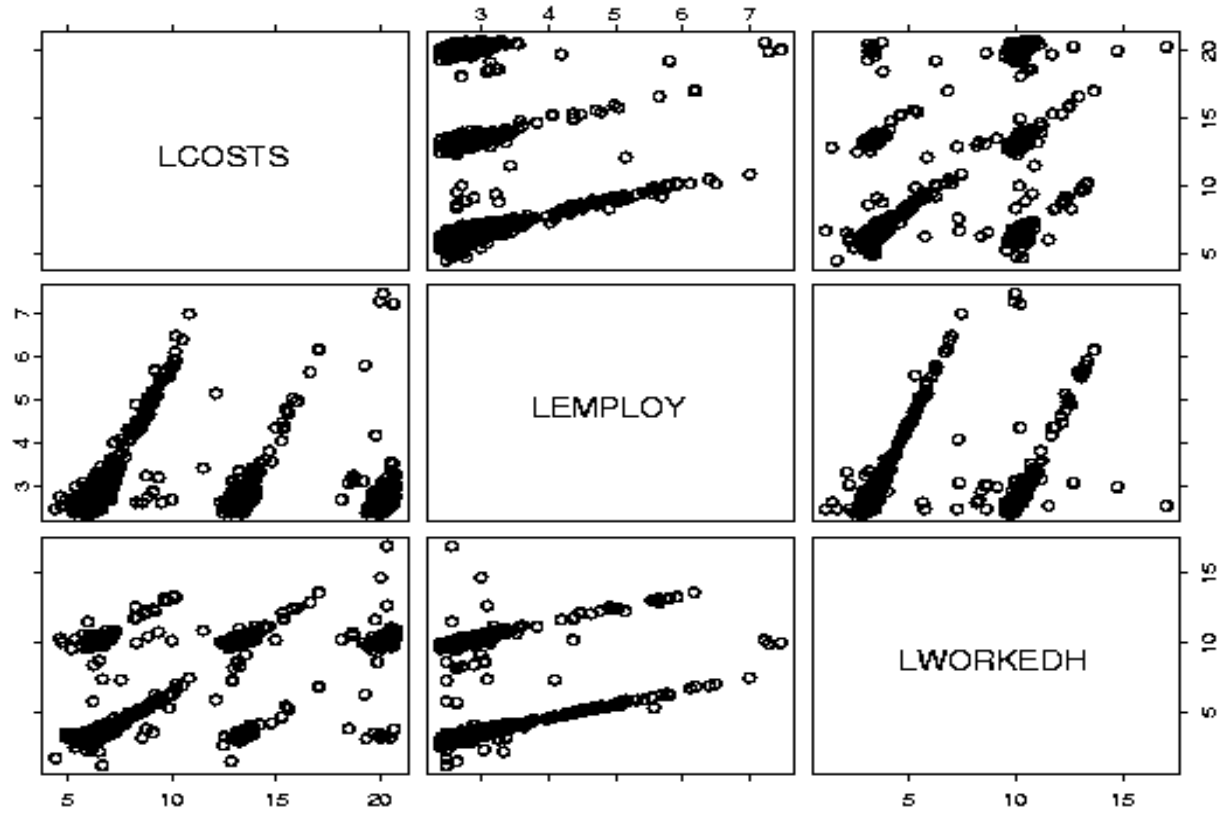
Using Mixture Modelling to Deal with Unity measure Error

- **Problem**: unity measure error times a costant factor (systematic error). E.g. costs × 1,000



- **Traditional approach** (max biv): EDA, ratio edits
- **Proposed approach**: mixture modelling

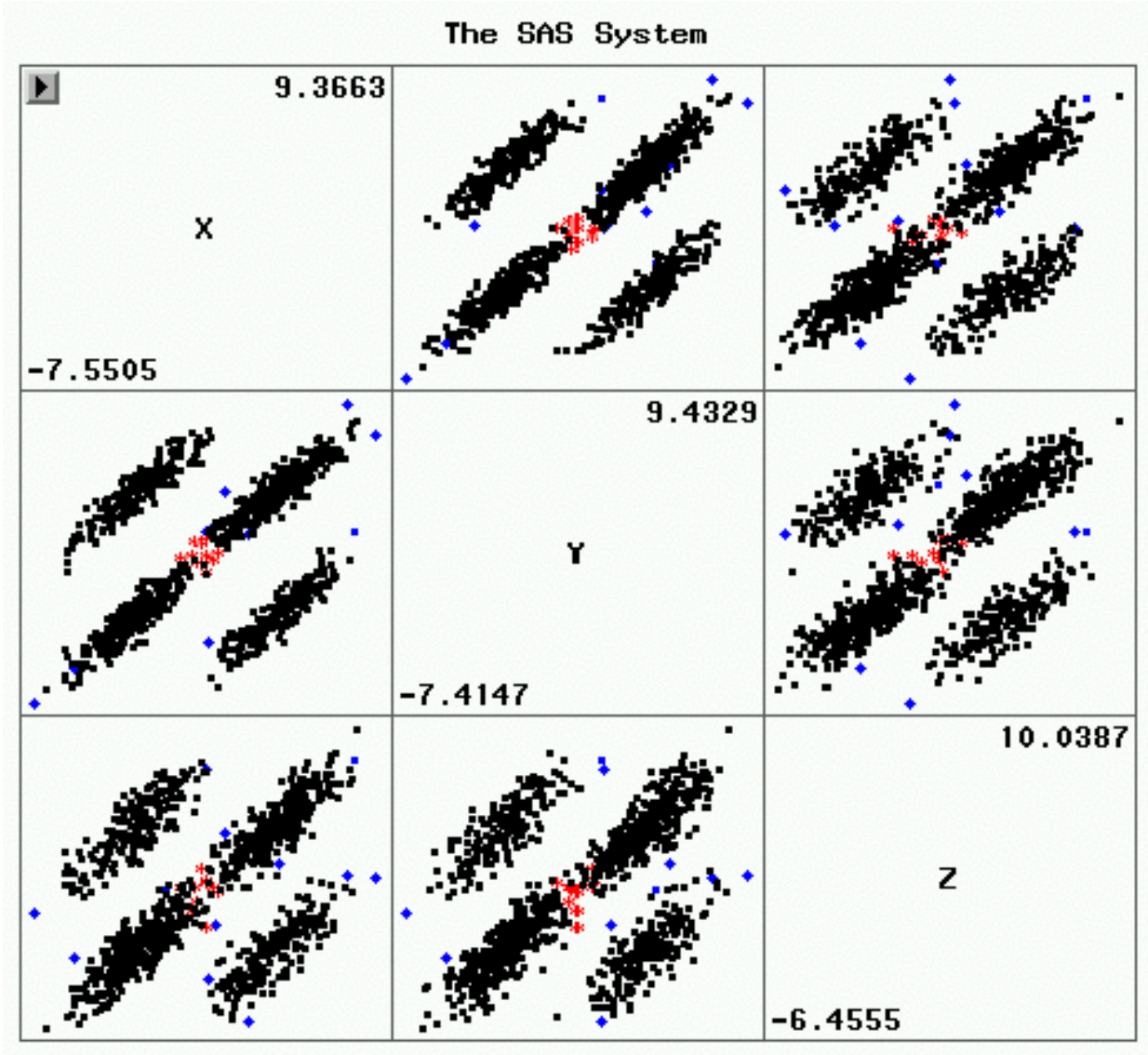
$$\sum_{t=1}^{h} \pi_t MN_t(\cdot; \mu_t, \Sigma)$$

– each mixture component corresponds to an error pattern G_t: e.g. $\mu_t=(E(X_1),E(X_2)+\ln(1{,}000),E(X_3))$

– *pr(j-th observation* $\in G_i | x^j ; \theta, \pi) =$ $\pi_i MN(\mathrm{x}^j; \mu_i, \Sigma) \Big/ \sum_{t=1}^{h} \pi_t MN(\mathrm{x}^j; \mu_t, \Sigma)$

Advantages

- Multivariate approach
- Provide diagnostic to combine in a common and coherent framework:
  - Automatic error detection
  - Selective/significance editing (uncertainty in the classification and/or atypicality w.r.t the model)


The SAS System
9.3663
X
-7.5505
9.4329
Y
-7.4147
10.0387
Z
-6.4555

Future work

- Build appropriate score function for selective editing on the basis of the model diagnostics
- Non-parametric approach