



Introduction à l'analyse multivariée analyse en composantes principales (ACP) et Partial Least-Square regression (PLS)

Mario Marchetti

Cerema Est – LMOPS/CentraleSupelec

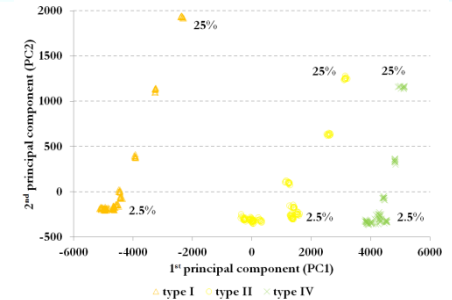


Fig.1: discrimination par ACP

PCA and PLS are statistical methods reducing data dimension through a projection onto a lower-dimensional space, when dealing with large datasets. In PCA, the variance-covariance matrix is used. Linear transformations of a group of correlated variables are obtained to reach optimal conditions. The transformed variables are uncorrelated, and resulting eigenvectors orthogonal. The physics generating the variations is substituted for a statistical approximation containing a linear combination of current physical factors. PLS consists in computing a calibration using a training dataset between some data and a parameter to forecast. A set of eigenvectors is searched for variables and another one for the parameter to forecast, to restore an optimum congruence between each variable factor and its corresponding in the least-squares sense. Several examples will be provided along with the theory.

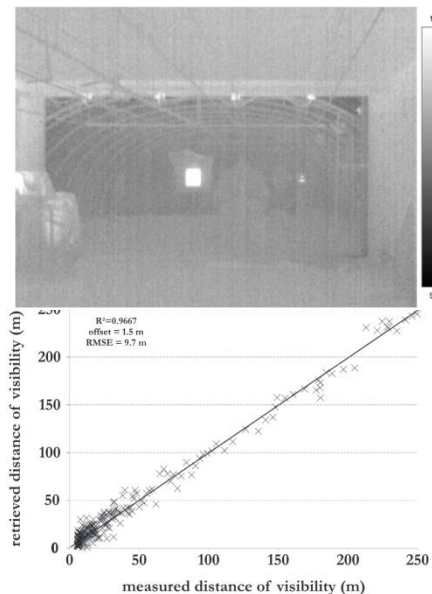


Fig.2: modèle PLS de visibilité dans le brouillard