**Reporting results**

This analyse takes into account the MOS obtained from three different experiments. Before to start any computing, the original MOS were executed to INSLA to put all experiments in a common scale. The design of the common set is of great importance for the mapping, and to represent precisely the relation between experiments. After all MOS values already are at the same scale, we are analysing them in two different cases:

CASE 1: It takes into account all results around of the experiments by trying to predict from Experiment 3 to Experiment 1 and 2 (and, the inverse, i.e., to predict from Experiment 1 and 2 to Experiment 3).

CASE 2: it takes into account only where Experiments have common setting artefacts (i.e., by discarding the same configuration shared by experiments).

After to make the necessary arrangement, a linear regression was executed by trying to explain the variability found, as following:

CASE 1a: linear regression was run by using (as DEPENDENT variable) and each artefacts as INDEPENDENT variables (i.e., packet loss, blockiness, and blurriness). The simple linear regression was used by trying to fit the results, given by: . The results have showed a Correlation=0,656 and an Average Error=0.137 when we tried to predict from Experiment 3 to Experiment 1. To the prediction from Experiment 3 to Experiment 2, the Correlation=0.897 and the Average Error=0.172. Figure 1a and 1b show the Observed vs Expected values to Experiment 1 and 2, respectively.

$$insla\_{MOS3}$$

$$\tilde{MOS\_{p}}=δ+α\*pck+β\*blo+γ\*blu$$



1. (b)

Figure 1 – Observed vs Expected values to Experiment 1 and 2, respectively

CASE 1b: a linear regression was run from a new variable, , created from and values (as DEPENDENT variable) and each artefacts as INDEPENDENT variables (i.e., packet loss, blockiness, and blurriness). The simple linear regression was used by trying to fit the results, given by: The results have showed a Correlation=0,912 and an Average Error=0.150 when we tried to predict from Experiment 1 and 2 to Experiment 3. Figure 2 show the Observed vs Expected values to Experiment 3.

$$insla\_{MOS12}$$

$$insla\_{MOS1}$$

$$insla\_{MOS2}$$

$$. \tilde{MOS\_{p}}=δ+α\*pck+β\*blo+γ\*blu$$



Figure 2 – Observed vs Expected values from inslaMOS12 to inslaMOS3

CASE 2: in this case, data were prepared to have only the common settings among experiments taking Experiment 3 as reference test. All were separated taking into account your artefacts (i.e., in Experiment 1 have just been selected the packet loss artefacts common in Experiment 3 and, Experiment 2 has two different artefacts, blockiness and blurriness common to Experiment 3. In this case, it was created two new columns taking into account just settings up where in common to Experiment 3). Each new column created has been treated as . Next, we have used the split sample validation concept to run a linear regression from (as DEPENDENT variable) and as INDEPENDENT variables (i.e., packet loss, blockiness, and blurriness MOS values, respectively). The simple linear regression was used by trying to fit the results, given by:

$$insla\_{MOS}$$

$$insla\_{pck}, insla\_{blo},insla\_{blu}$$

$$insla\_{MOS3}$$

$$insla\_{pck}, insla\_{blo},insla\_{blu}$$

. The first group (split=0) showed a Correlation=0,825 and an Average Error=0.080. Second group (split=1) have showed a Correlation=0,828 and an Average Error=0,076, when we have tried to predict from to Experiment 3. Figure 3a and 3b show the Observed vs Expected values to the Split=0 and Split=1 groups, respectively.

$$\tilde{MOS\_{p}}=δ+α\*insla\_{pck}+β\*insla\_{blo}+γ\*insla\_{blu}$$

$$insla\_{pck}, insla\_{blo},insla\_{blu}$$



(a) (b)

Figure 3 – Observed vs Expected values of inslaMOS3 from split sample validation