Multilinear Regression for Gamma Expansion of Overexposed Content

Technical Report RR-03-11
Universidad de Zaragoza

Belen Masia
Universidad de Zaragoza

Diego Gutierrez
Universidad de Zaragoza

July 11, 2011

1 Introduction

Reverse tone mapping techniques aim at increasing the dynamic range of legacy low dynamic range images, generally for viewing on high dynamic range displays. Almost all existing reverse tone mapping operators (rTMOs) follow the same strategy: they first compute some sort of *expand map* based on detecting high-luminance areas of the LDR input images; pixels included in such expand map are then aggressively expanded, while the rest are either left unchanged or only slightly expanded. While this strategy works well for correctly exposed images (where typically only a small percentage of pixels are saturated), its performance decreases significantly when the input images are overexposed to begin with, as was shown by Masia et al. [4].

In their work, the authors show that, for overexposed input content, a global rTMO based on a simple gamma expansion outperforms more sophisticated algorithms, across a range of different overexposure levels. This finding should help guide the design of future reverse tone mapping operators. The authors additionally propose an easy method to obtain a suitable γ value for each image, based on computing its *key* value, which is an indicator of whether the scene is subjectively dark or light.

This work extends the original publication by Masia et al. by further investigating other image metrics and statistics, which can be used to compute suitable γ values for the low dynamic range input images. We propose new, robust models that are still straightforward to compute, while showing

a higher correlation with the data gathered in the user studies [4]. Additionally, we note that Section 5 in the original paper incorrectly reported that a linear regression of the form $\gamma = 10.44k - 6.282$ was used.

Thus, given the new statistical analysis in this work, different parametric regressions are obtained together with various metrics, to assess how well they fit the data. Both single variable regressions and more complex multilinear least squares fittings are computed. Data fitting by means of linear least squares is performed for different subsets of image statistics, always seeking to obtain the best trade-off between accuracy and simplicity of the model.

2 Manually adjusted γ values

Table 1 includes the γ values for the image database which were manually adjusted to obtain the subjective best depiction. Columns 1 to 4 indicate increasing exposure, while rows correspond to the different scenes captured. Figures 2 and 3 in the original publication [4] show the nine different scenes plus the complete bracketed sequence for two of them, respectively. Within each bracketed sequence, images present an increasing degree of overexposure.

Table 1: Manually adjusted γ values

Scene	1	2	3	4
Building	1.22	1.5	1.75	2.6
Lake	1.1	1.2	1.5	2.25
Sunset	1.1	1.35	1.4	1.75
Graffiti	1.2	1.35	1.5	1.75
Strawberries	1.22	1.35	1.55	1.9

3 Image Statistics

We compute below a series of statistics. For all of them, luminance is obtained from sRGB linearized values [5]:

$$L = Y = 0.2126R + 0.7152G + 0.0722B, \tag{1}$$

where L is thus normalized to [0..1].

These statistics include both the arithmetic and the geometric mean luminance (referred to as L_{avg} and L_H, respectively). The arithmetic mean is simply obtained by averaging the luminance value of all pixels ($L_{avg} = 1/N \sum_{i=1}^{N} L(i)$, with N being the total number of pixels in the image); the geometric mean, known to reduce the contribution of outliers, is obtained as follows [2]:

$$L_H = exp\left(\frac{1}{N}\sum_{i=1}^{N} log(L(i)+\varepsilon)\right), \tag{2}$$

where ε is a very small positive number to prevent singularities in black pixels. We additionally compute the logarithm of this quantity, simply $logL_H$.

The key of the images is also obtained, using the following equation [1]:

$$k = \frac{logL_H - logL_{min}}{logL_{max} - logL_{min}}. \tag{3}$$

In this equation L_{max} and L_{min} are the maximum and minimum luminance values, respectively, once a percentage of outlier pixels (both on the dark and bright sides) has been eliminated. We calculate two key values, k_5 and k_1, considering 5% or 1% of the pixels as outliers, respectively.

Additionally, both the median, L_{med}, and a series of central moments, are computed for the luminance of the images. These include variance V_L (and standard deviation σ_L), skewness ($skew_L$) and kurtosis ($kurt_L$). Finally, we compute the percentage of overexposed pixels for each of the images, defining overexposed pixels as those with $L \cdot 255 \geq 254$; we will refer to it as p_{ov}.

Table 2 includes the values obtained for each of the aforementioned statistics for the images of our dataset. Please note that the values for the key of the image slightly differ from those appearing in the original paper [4]. This values shown in Table 2 are the ones used for the regressions explored in the following sections.

4 Fitting the data: multilinear regression

The straightforward initial step is to obtain a regression between γ and one –and only one– of the image statistics previously computed, in order to keep the model as simple as possible. Typically this would be a linear regression, as suggested in [4], although in certain specific cases a better fitting is obtained with an exponential or a potential regression. However, none of the regressions obtained with a single variable had an R^2 value greater than 0.7, and were therefore discarded in search of a better fit.

Table 2: Statistics for the images in our dataset.

Image	γ	L_{avg}	L_H	$logL_H$	k_5	k_1	L_{med}	V_L	σ_L	$skew_L$	$kurt_L$	p_{ov}
Building01	1.22	0.3493	0.1182	-2.1352	0.5743	0.6019	0.0762	0.1357	0.3684	0.5261	1.6310	5.9764
Building02	1.5	0.4853	0.2485	-1.3922	0.6472	0.6775	0.2076	0.1774	0.4212	0.1879	1.1346	20.0422
Building03	1.75	0.5792	0.4052	-0.9033	0.6865	0.7265	0.4176	0.1527	0.3908	0.0275	1.1750	40.1717
Building04	2.6	0.7105	0.6196	-0.4787	0.7392	0.7806	0.7399	0.0912	0.3020	-0.4495	1.7895	44.3981
Lake01	1.1	0.1188	0.0338	-3.3881	0.4963	0.5612	0.0248	0.0256	0.1601	1.1686	2.6714	0.0003
Lake02	1.2	0.1662	0.0570	-2.8651	0.5151	0.5721	0.0473	0.0450	0.2121	1.1497	2.6493	0.0017
Lake03	1.5	0.3689	0.2020	-1.5996	0.5545	0.6418	0.1827	0.1316	0.3628	0.8654	2.1098	17.8714
Lake04	2.25	0.4977	0.3613	-1.0182	0.5975	0.6874	0.3668	0.1225	0.3500	0.4138	1.5494	22.8514
Sunset01	1.1	0.2088	0.0719	-2.6319	0.4857	0.5622	0.0784	0.0786	0.2803	1.6711	4.7589	4.0931
Sunset02	1.35	0.2633	0.1143	-2.1688	0.5252	0.6170	0.1361	0.0888	0.2980	1.3072	3.5468	5.0731
Sunset03	1.4	0.3930	0.2259	-1.4875	0.5884	0.6907	0.2675	0.1152	0.3394	0.6207	1.9193	9.4525
Sunset04	1.75	0.6633	0.5505	-0.5969	0.6891	0.8009	0.7168	0.1025	0.3202	-0.3523	1.6076	29.5133
Graffiti01	1.2	0.2891	0.1568	-1.8525	0.6213	0.6726	0.2744	0.0557	0.2361	0.3878	1.8997	0.0039
Graffiti02	1.35	0.5020	0.3335	-1.0981	0.6659	0.7405	0.5732	0.1110	0.3332	-0.0818	1.4351	1.0967
Graffiti03	1.5	0.6796	0.5449	-0.6071	0.7201	0.8074	0.8761	0.1158	0.3403	-0.5197	1.5827	21.1744
Graffiti04	1.75	0.8091	0.7415	-0.2991	0.7761	0.8591	0.9949	0.0692	0.2631	-1.0777	2.8062	51.2861
Strawberries01	1.22	0.1718	0.0954	-2.3501	0.5646	0.6074	0.1075	0.0283	0.1681	1.2238	3.6194	0.0000
Strawberries02	1.35	0.3381	0.2240	-1.4962	0.5713	0.6335	0.2544	0.0729	0.2700	0.7483	2.3469	0.0781
Strawberries03	1.55	0.5304	0.4098	-0.8921	0.6235	0.6884	0.5048	0.1037	0.3220	0.1416	1.5362	8.9808
Strawberries04	1.9	0.6961	0.6046	-0.5033	0.6942	0.7524	0.8147	0.0934	0.3057	-0.4674	1.6204	26.5756

Since none of the single variable regressions provided a significantly high R^2 value we tried fitting the data with multilinear regressions, that is, linear regressions with multiple variables as predictors. Restricting ourselves to linear regressions was decided to keep the model as simple as possible; if a good model could not be found assuming a linear relationship, we would move on to more complex fittings. We initially used ordinary least squares to do the fittings. This implies a series of assumptions over the errors, mainly that they are normally distributed, with constant variance, and independent of each other. It also implies that the independent variables are free of error, or that their error is insignificant compared to the error of the dependent variable. We tested and analyzed several different fittings, varying the subset of image statistics that constituted the independent variables.

Once the type of model (i.e. linear) has been chosen, the problem which arises when working with multiple predictors is knowing which of the possible predictors (i.e. the independent variables, in our case the calculated image statistics) should be included in the model and which should be left out.

The way in which we deal with this is performing F-tests over the possible models. Computing the R^2 value or another goodness of fit metric and comparing their values for both models is typically not enough. The reason for this is that given two models, A and B, with p_A and p_B terms, respectively, if $p_A > p_B$, model A will always fit the data at least as good as model B. Thus, what has to be found out is if the addition of that extra parameter(s) to model A gives a *significantly* better fitting; as mentioned, we make use of F-statistics to assess that. Annex A describes the use of F-tests in the construction of multiple variable models.

4.1 Building multiple variable models

A stepwise regression is used to build the possible multilinear models [6]. The idea is to start from an initial model, typically with the single variable showing the highest correlation with the data. Then, the steps described below are performed iteratively until no more variables can be added.

- **Step 1:** Test all variables independently with partial F-statistics. The variable with the lowest p-value is added to the model (as long as the p-value is lower than the significance level, p-value $< \alpha$). This step is repeated until no more variables can be added.
- **Step 2:** Since the addition of a new variable can cause a previously introduced value to be non-significant, a partial F-statistic is computed

for each variable currently in the model as if it were the one introduced in the last place. The one with the largest p-value is removed (as long as the p-value is larger than twice the significance level, p-value $> 2\alpha$), and a new iteration begins with Step 1.

The process ends when, in Step 2, no variables are removed because there are no variables whose p-value is larger than the significance level (which implies that no variables can be added either). Please note that even though some combinations of variables may not have much sense (e.g. having k_5 and k_1 both included in a regression) we initially make exhaustive tests without taking into consideration these constraints, which will be evaluated over the final model selected.

It is clear from the procedure above that the initial model selected and the order in which variables are selected to be introduced or removed from the model determines the final model reached. For this reason, we repeated the algorithm for stepwise regression starting from every possible single variable initial model. This yields different final models, for which a series of metrics are computed in order to evaluate the accuracy of the fitting. In particular we compute the RMSE and the *overall* F-statistic for each model obtained (see Annex B for a definition of these parameters). Based on them we select the best model, which yields the following equation for the regression:

$$\gamma = 3.8872 + 0.3752 log L_H - 2.9941 k_1 + 0.0160 p_{ov}. \quad (4)$$

Figure 1 shows the observed γ values against the γ values predicted by this model. Additionally, it compiles different metrics assessing the accuracy of the previous regression, including R^2 and $\tilde{R}^2$ (see Annex B for a definition of $\tilde{R}^2$) as well as the already mentioned RMSE and F-statistic with its associated p-value.

There is another possible model resulting from the stepwise fitting process which involves L_{med} and L_H and yields a lower RMSE than the model given by Equation 4. However, when using this model three observations are classified as outliers, which we found unjustified given the slight increase in performance.

4.2 Checking for outliers

The procedure described in Section 4.1 yields several different models. However, when performing an outlier analysis over the residuals of these models, there are two observations which systematically appear as outliers in the fitted models (the 95% confidence interval for the mean of their residuals does

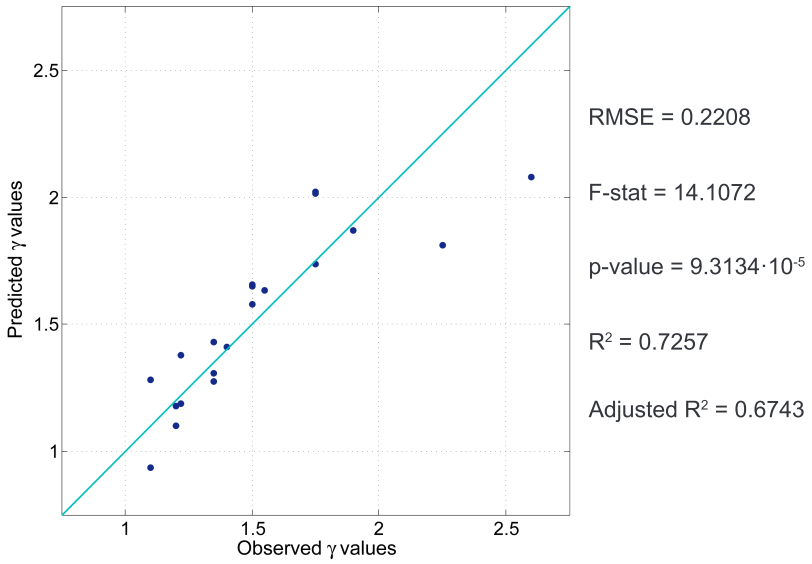

Figure 1: Predictive accuracy of the regression shown in Equation 4. The x-axis shows observed γ values, while the y-axis depicts the values predicted by the regression. The cyan line shows the quadrant bisection (i.e. y=x).

not include zero). We therefore chose to remove these two values (*Building04* and *Lake04*) from the data and re-do the multilinear regression fitting.

A detailed summary of the different models obtained using stepwise regression on the data without outliers is shown in Table 3. The first column shows the variable included in the initial model, the second column the variables included in the final model, and the third column their associated p-values. The p-value is that of the variable in the final model (not necessarily equal to the p-value at the point in which it was introduced). Finally, the last columns show the RMSE and the overall F-statistic with its associated p-value for that particular model, as a measure of the accuracy of the fitting.

4.3 Selecting the best multilinear model

According to Table 3, the model which yields the best fit to the data –i.e. the one with the lowest RMSE, and with the lowest p-value in the overall F-test– is the one given by the following equation:

$$\gamma = 0.9855 + 2.8972L_H - 0.8232L_{med} + 0.2734skew_L - 0.0898kurt_L \quad (5)$$

Table 3: Different models obtained by stepwise regression.

Variable in initial model	Variables in final model	p-value	RMSE	F (p-value·10^5)
L_{avg}	L_{avg}	$6.3445 \cdot 10^{-7}$	0.1121	62.6831 (0.0635)
L_H	L_H	$9.3551 \cdot 10^{-6}$	0.0810	67.9410 (0.0030)
	L_{med}	$7.9268 \cdot 10^{-4}$		
$logL_H$	$logL_H$	$1.1925 \cdot 10^{-3}$	0.0933	49.3436 (0.0253)
	p_{ov}	$5.3205 \cdot 10^{-3}$		
k_5	L_{avg}	$6.3445 \cdot 10^{-7}$	0.1121	62.6831 (0.0635)
k_1	L_{avg}	$6.3445 \cdot 10^{-7}$	0.1121	62.6831 (0.0635)
L_{med}	L_H	$9.3551 \cdot 10^{-6}$	0.0810	67.9410 (0.0030)
	L_{med}	$7.9268 \cdot 10^{-4}$		
V_L	L_H	$9.3551 \cdot 10^{-6}$	0.0810	67.9410 (0.0030)
	L_{med}	$7.9268 \cdot 10^{-4}$		
σ_L	L_H	$9.3551 \cdot 10^{-6}$	0.0810	67.9410 (0.0030)
	L_{med}	$7.9268 \cdot 10^{-4}$		
$skew_L$	L_{avg}	$6.3445 \cdot 10^{-7}$	0.1121	62.6831 (0.0635)
$kurt_L$	L_H	$2.7642 \cdot 10^{-6}$	0.0664	52.8833 (0.0065)
	L_{med}	$6.7751 \cdot 10^{-3}$		
	$skew_L$	$1.6002 \cdot 10^{-2}$		
	$kurt_L$	$1.0849 \cdot 10^{-2}$		
p_{ov}	$logL_H$	$1.1925 \cdot 10^{-3}$	0.0933	49.3436 (0.0253)
	p_{ov}	$5.3205 \cdot 10^{-3}$		
None	L_{avg}	$6.3445 \cdot 10^{-7}$	0.1121	62.6831 (0.0635)

The goodness of fit of this regression can be measured through different metrics. Some of them, such as the RMSE and the F-statistic, have already been included in Table 3. Nevertheless, Figure 2 compiles them together with other metrics, namely R^2 and $\tilde{R}^2$. To further illustrate the predictive accuracy of our model, Figure 2 shows the observed γ values in the x-axis, against the γ values predicted by our model.

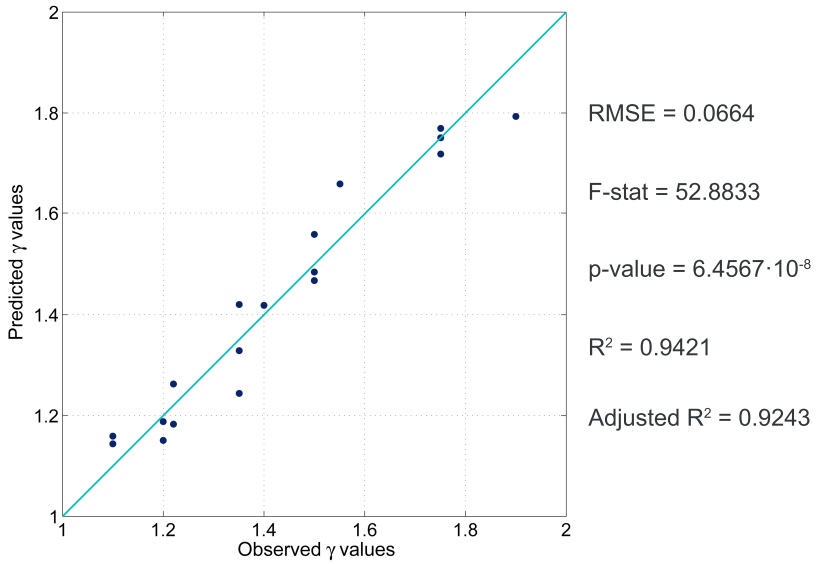

Figure 2: Predictive accuracy of the regression shown in Equation 5. The x-axis shows observed γ values, while the y-axis depicts the values predicted by the regression. The cyan line shows the quadrant bisection.

4.4 Robust regression

An alternative to eliminating the observations regarded as outliers as done in Sections 4.2 and 4.3 would be to retain all the observed data but weight their influence when computing the regression. To do this, we perform a new regression with the same predictors as in Equation 4 using iteratively reweighted least squares. The weight function used is a bisquare function. The new regression is thus given by the following equation:

$$\gamma = 2.4379 + 0.2319 logL_H - 1.1228 k_1 + 0.0085 p_{ov}. \quad (6)$$

Figure 3 shows the predictive accuracy of the model obtained by robust regression compared to Ordinary Least Squares, both with all the observed

data (*left*) and after outlier removal (*right*). Additionally, if we compute a robust RMSE estimate for this last regression [3], we obtain an estimate of 0.0962 (while estimates for the previous ones, OLS with and without oultiers, were 0.2208 and 0.0664 respectively).

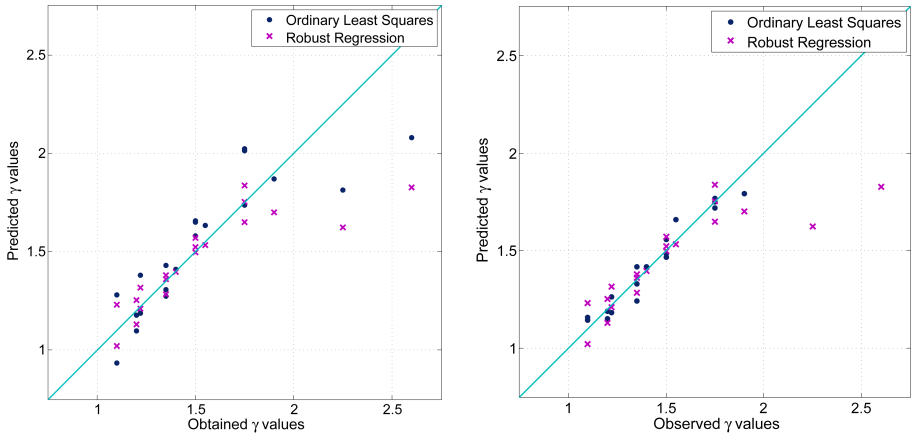

Figure 3: Predictive accuracy of the model obtained by robust regression against the ones obtained by ordinary least squares, with all the observed data (*left*) and after outlier removal (*right*). The abscissa show observed γ values, while the y-axes depict the values predicted by the regression. The cyan lines mark the quadrant bisection.

5 Acknowledgements

We thank Rafael Pacheco Kovaleski and Francesco Banterle for their insightful comments on our original paper [4]. This research has been partially funded by a Marie Curie grant from the 7th Framework Programme (251415), the Spanish Ministry of Science and Technology (TIN2010-21543) and the Gobierno de Aragón (OTRI 2009/0411 and CTPP05/09). Belen Masia is supported by an FPU grant from the Spanish Ministry of Education.

References

[1] A. O. Akyuz and E. Reinhard. Color Appearance in High Dynamic Range Imaging. *Journal of Electronic Imaging*, 15(3), 2006.

[2] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers. *Advanced High Dynamic Range Imaging: Theory and Practice*. A.K. Peters, Ltd., 2011.

[3] W. Dumouchel and F. O'Brien. *Integrating a robust option into a multiple regression computing environment*, pages 41–48. Springer-Verlag New York, Inc., New York, NY, USA, 1991.

[4] B. Masia, S. Agustin, R. W. Fleming, O. Sorkine, and D. Gutierrez. Evaluation of reverse tone mapping through varying exposure conditions. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia)*, 28(5):160:1–160:8, 2009.

[5] E. Reinhard, E. A. Kahn, A. O. Akyuz, and G. M. Johnson. *Color Imaging: Fundamentals and Applications*. A.K. Peters, 2008.

[6] J. H. Steiger. Introduction to multiple regression. `http://www.statpower.net/Content/312/Lecture%20Slides/MultipleRegressionIntro.pdf`. [Online, last accessed 30-June-2011].

A F-tests for assessing the appropriateness of adding new predictors to a model

An F-test is typically performed to decide whether or not a certain null hypothesis can be rejected. To do this, a test statistic (the F-statistic) is needed which under the null hypothesis follows an F-distribution. In our case, the null hypothesis is that, given two models, A and B, with a number of predictors p_A and p_B ($p_A > p_B$), the two models fit equally well the data. The F-statistic is then given by:

$$F_{p_A-p_B,n-p_A} = \frac{(SS_B - SS_A)/(p_A - p_B)}{SS_A/(n-p_A)}, \tag{7}$$

where SS_i, $i = \{A, B\}$, is the sum of squared residuals of model i, and n is the number of data values [6]. It must be noted that in Equation 7, and throughout the document, p_i as a measure of the number of terms in the regression includes the constant term (i.e. the intercept).

For the particular case of creating model A by adding one variable to a model B that has p terms, and expressing the formula in terms of R^2, the F-statistic becomes:

$$F_{1,n-p-1} = \frac{R_A^2 - R_B^2}{(1-R_A^2)/(n-p-1)} \tag{8}$$

As it is well known, given a value for F in an F-test, the p-value is the probability of obtaining a value as extreme as the F obtained, assuming that the null hypothesis is true. As a consequence, the null hypothesis is typically rejected if the p-value is lower than the significance level alpha (which, in this work, will have the usual value of $\alpha = 0.05$).

B Goodness of fit in multilinear regressions

This Annex includes the description of a series of metrics which are typically used in regression analysis to measure the accuracy of the fitting of a certain model.

RMSE. For a multilinear regression, RMSE is computed as shown in Equation 9, where Y_i are the observed data (i.e. the given γ values) and $\hat{Y}_i$ the data predicted by the model.

$$RMSE = \sqrt{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2/(n-p)}, \tag{9}$$

where, n is the data size and p the number of terms in the regression. Please recall that in this formulation the intercept is included in p. This metric provides an intuition on the error we would incur in when using a certain regression to estimate the value of a variable.

Overall F-statistic. The *overall* F-statistic is simply an F-test in which the null hypothesis is that the data can be explained by a constant (which would be the mean of the observed data), versus the hypothesis that the data can be explained by the selected model. Therefore, a high F-statistic and, specially, a low associated p-value indicate that the hypothesis that our model explains the data (vs. the hypothesis that a constant explains them) is clearly correct.

R^2 and adjusted R^2. Typically used to assess how well the values predicted by a model will adjust to the real values, in the case of linear regressions R^2 is simply the square of the correlation coefficient between the observed and the predicted data.

However, in the case of multilinear regression, the R^2 value will always increase as new variables are added to the model. For this reason sometimes the *adjusted* R^2 is used, which corrects for the number of explanatory

variables in the model. As a result, the adjusted R^2 value will only increase if the new term improves the regression more than would be expected by chance. The adjusted R^2 value is usually denoted by $\tilde{R}^2$ and computed as follows:

$$\tilde{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-p} \tag{10}$$

where, again, n is the data size and p the number of terms in the regression. Please recall that in this formulation the intercept is included in p. It is well-known that the higher the R^2 and the adjusted R^2 values, the higher the correlation between the values predicted by the model and the values actually observed.