Shape analysis of concrete aggregates for statistical quality modeling *

Ilkka Kalliomäki, Aki Vehtari and Jouko Lampinen
Laboratory of Computational Engineering,
Helsinki University of Technology,
P.O.Box 9203, FIN-02015 HUT, Finland.

February 21, 2005

Abstract

We analyze images of concrete aggregate material and design computational measures for angularity, surface texture and shape of the particles. The features obtained with image analysis are used as additional explanatory variables in statistical models for predicting the quality properties of concrete. Model comparison confirms that the image analysis features improve predictions.

1 Introduction

The goal of the presented work was to develop models for predicting how the aggregate characteristics affect quality properties of concrete, as a part of a large quality control program of the industrial partner. The need for such models is significant as conventional concrete aggregate supplies are becoming depleted, and environmental aspects limit the use of existing sources. The quality variables included, for example, compression strengths and densities for 1, 28 and 91 days after casting, and bleeding (water extraction), flow value, slump and air percentage, which measure the properties of fresh concrete. These quality measurements depend on the properties of the aggregates (stone material), additives, and the amount of cement and water used. Aggregate properties are, for example, natural or crushed, size and shape distributions of the grains, and mineralogical composition.

In this paper we describe the property measurements obtained using image analysis, and demonstrate that better predictions are achieved with these image analysis measurements, even when the effect of finite data set size is taken into

*This is an electronic preprint version of the article which is to appear in Machine Vision and Applications, Springer, 2005. The original publication is available at www.springerlink.com, DOI: 10.1007/s00138-004-0172-3 URL: http://www.springerlink.com/link.asp?id=fycb2jmy4guc29e5

account. In the study, we had a total of 27 explanatory variables: 3 describing the basic recipe, 8 from image analysis, and 16 other variables. 215 samples designed to cover the practical range of the variables were collected by the concrete manufacturing company. The details of the problem and the conclusions made by the concrete expert are described in Ref. (Järvenpää, 2001).

2 Image analysis

Images of the aggregate material were obtained with a 1200 DPI flatbed scanner against a matte black velvet background. Image resolution is approximately 4000 by 3000 pixels. Each image contains up to several hundred particles, which were manually arranged not to overlap or touch each other. Our explanatory variables obtained with image analysis include angularity, surface texture, elongation and flakiness. These relatively easily computable attributes of the particles are inspired by physical measurements, traditionally determined using measuring gauges and subjective evaluation. We analyzed a total of seventeen aggregate types with three different size fractions each. Previously aggregate shape has been analyzed using image processing methods in predicting asphalt pavement performance (Masad and Button, 2000), where the very small particle size required the use of a microscope in image acquisition and a considerably smaller number of particles was analyzed in total.

3 Segmentation and pre-processing

We use a hard segmentation strategy and estimate the best threshold for an image by first finding the local minimum in the gray level histogram between dark background pixels and lighter particle pixels, multiplied by a hand-adjusted constant so that a desired segmentation result was obtained in a number of test images.

The thresholded image is filtered with a 3-by-3 pixel median filter and holes inside particles are filled. It is not crucial for the application that all particles are detected perfectly, while incorrectly segmented particles can contaminate the data quite badly. For this reason, as a final preprocessing step, the data is automatically censored by excluding from further analysis those particles which are highly anomalous in their size or compactness values. This could be avoided with a more systematic imaging setup, for example by organizing the particles in a grid in which it is guaranteed that no two particles touch each other. Fig. 1 shows a typical input image and the convex hulls of recognized particles overlayed on top of it.

3.1 Angularity

We design a novel measure for the angularity of the particle using morphologic operations. The method is based on the erosion operation using a circular structur-

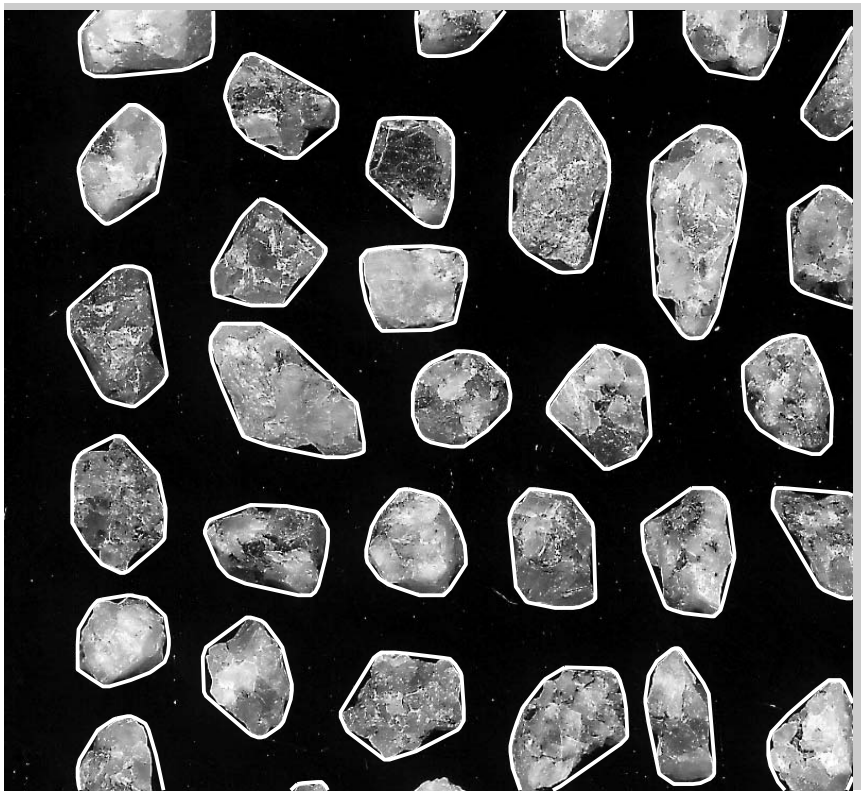


Figure 1: Recognized particles outlined in input image.

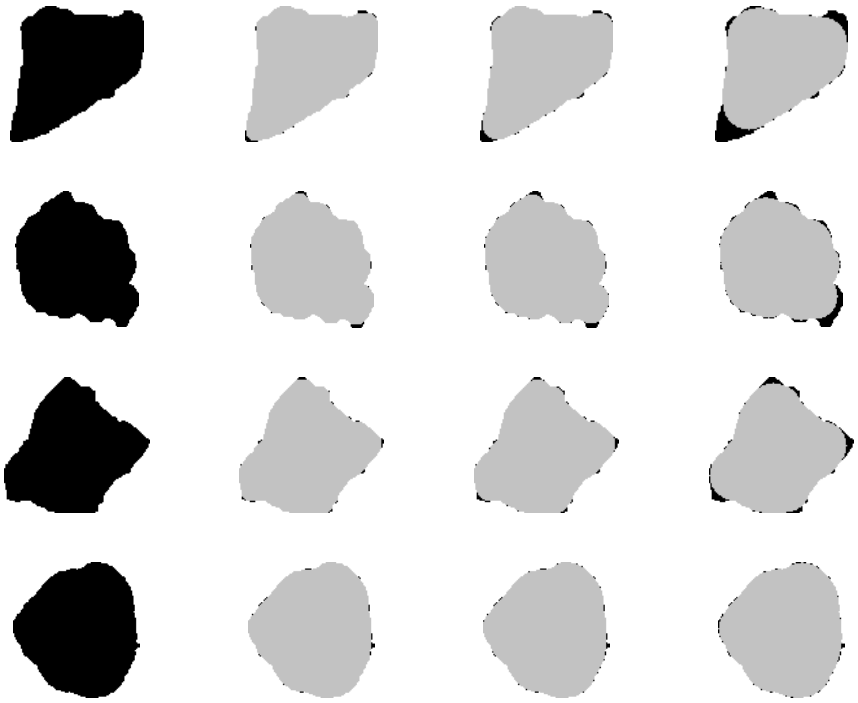


Figure 2: Morphological erosion of differently shaped particles. The leftmost column shows the original particle shapes. The three following columns correspond to successively increasing mask diameters of 7, 10 and 18 pixels, showing the eroded pixels in black. The rightmost column has a larger mask than what was used in the analysis for illustrative purposes.

ing element (or mask) of increasing diameter. The number of eroded pixels as a function of mask diameter is denoted the *morphologic spectrum* of the image.

Using the morphologic spectrum, we define the angularity of the particle as

$$a = \sum_{i=1}^{5} \frac{n_i}{d_i^2}, \tag{1}$$

where i refers to the mask index, n_i is the number of eroded pixels with mask i, and d_i is the mask diameter. Circular masks of diameters $d \in \{3, 5, 7, 10, 12\}$ pixels were used. The scaling factor $1/d_i^2$ gives roughly equal importance to different mask sizes. Typically recently crushed rock has distinctly sharp edges, whereas natural gravel, eroded by the ice age, has round edges. Sharp corners become rounded in erosion, and a large number of pixels become eroded especially in triangular particles. In contrast, very little erosion occurs on round, smooth particles. Figure 2 shows four particles, in decreasing order of angularity from top to bottom, and their eroded images with masks size increasing from left to right.

3.2 Surface texture

In addition to the silhouette of the particle, its texture inside the borderline contains information about the roughness of the particle surface. We use the energy contained in a number of frequency bands (Tuceryan and Jain, 1993) as a measure for roughness. First we compute the 2D Fourier transform of a square texture patch which is contained completely inside the particle. The Fourier transform of the texture patch, normalized in size by continuing it periodically in the spatial plane, is multiplied with radial Gaussian filters

$$G_i(\omega_x, \omega_y) = \exp\left(-\frac{\left(\sqrt{\omega_x^2 + \omega_y^2} - c_i\right)^2}{2\sigma_i^2}\right) \tag{2}$$

where c_i is the filter center frequency and σ_i the filter bandwidth. A downside of the periodical continuation of the texture patches is that it can introduce artificial edges especially in smooth textures. There is often considerable shading in the images of smooth stones, and artificial edges can have a significant effect to their energy content. However, in general the texture features are not severely contaminated with the edge effects, because their energy is small compared to the total variation in surface texture energy between different aggregate types.

We define the texture energy measure vector s with elements s_i as

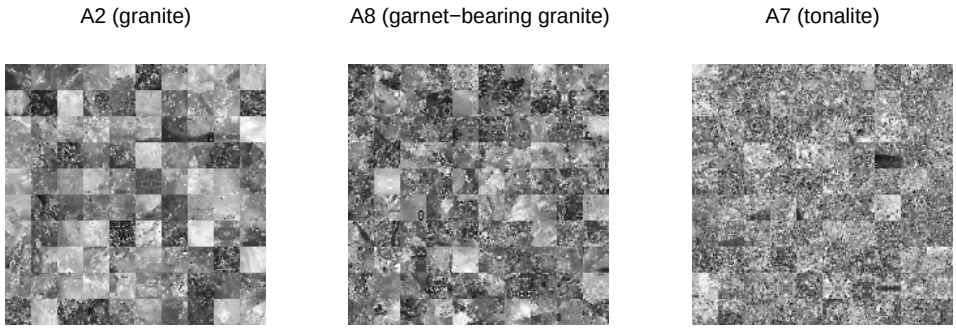
$$s_i = \sum_{\omega_x} \sum_{\omega_y} |G_i(\omega_x, \omega_y) T(\omega_x, \omega_y)|^2, \tag{3}$$

where $T(\omega_x, \omega_y)$ is the Fourier transform of the periodically continued texture patch $t(x, y)$. The summation is performed over the whole frequency space. We used five differently tuned filters G_i with octave spacing in their center frequencies $c_i \in \{1/2, 1/4, 1/8, 1/16, 1/32\}$ (in units of cycles/pixels) and bandwidths σ_i.

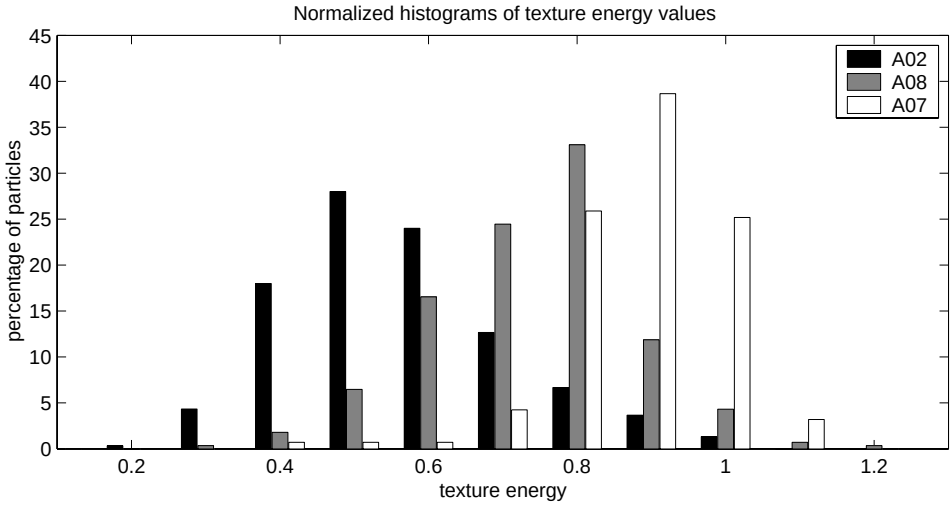
Figure 3 shows texture samples obtained from three different types of aggregate and energy histograms of their textures in the second highest frequency band (wavelength of four pixels). The variances of the distributions are approximately equal, but aggregate A7 has clearly the largest amount of energy in average, compared to A8 and A2. Relative differences in texture energy between aggregate types are largest in the highest frequency bands. All five frequency bands were used as features. Note that Figure 3b is not a distribution of overall texture energy in a single texture sample, but rather the histogram of energy in a single frequency band across all 100 texture samples in Figure 3a.

3.3 Elongation and flakiness

Measures for overall roundness of the stones can be defined by taking into account the way in which differently shaped stones tend to lie on the imaging surface. We define two dimensionless measures for particle shape which are both based on the



(a) Texture samples



(b) Corresponding distributions of texture energies in a single frequency band

Figure 3: Texture samples (top) and their energy histograms (bottom) of three different aggregate types, one hundred particles each. The texture energy of smooth granite (A2) at a high frequency band is significantly lower than that of the rugged tonalite (A7).

number of pixels in each particle. Additionally, flakiness depends also on physical volume measurements.

Elongation is defined as

$$e = \frac{\varphi_{2D}}{\varphi_{sphere}} = \frac{2\sqrt{A/\pi}}{\varphi_{sphere}} \propto \sqrt{A}. \tag{4}$$

In this expression $\varphi_{2D} = 2\sqrt{A/\pi}$ is the diameter of a circle with equal area A on the image plane. φ_{sphere} is the maximum diameter of a 3D sphere which can pass through the sieve. It is constant for a given aggregate size fraction, and makes the elongation number dimensionless. Thus elongation is directly proportional to the square root of number of pixels in the image of the particle. Elongated stones can pass the sieve longways, but will lie sideways on the flat imaging surface, and will have large equivalent circle diameter. Approximately spherical pebbles have the lowest and stick-shaped stones the highest elongation numbers.

Flakiness is defined as

$$f = \frac{\varphi_{2D}}{\varphi_{3D}} = \frac{2\sqrt{A/\pi}}{2\sqrt[3]{V/(4\pi/3)}} \propto \frac{\sqrt{A}}{\sqrt[3]{V}}, \tag{5}$$

where $\varphi_{3D} = 2\sqrt[3]{V/(4\pi/3)}$ is the diameter of a sphere with equivalent average volume. The volume is computed by counting the number of particles and physically measuring the total volume of each aggregate sample individually. Flakiness is proportional to square root of the image area of the particle and inversely proportional to the cubic root of average volume of the aggregate sample. The flakiness number is largest for flat and thin, coin-like stones and smallest for spherical pebbles.

The particle count was computed only partially automatically. Because some of the images contained particles that touch each other, the automatically obtained particle counts were cross-checked manually. The use of a special purpose particle counter would eliminate the need for manual counting.

4 Regression model

Based on the expert knowledge and preliminary analysis it was known that there were nonlinear effects, strong interactions and dependencies in the covariates. A non-linear nonparametric Gaussian process model is useful in such situations because it can handle these interactions implicitly. The Gaussian process model is a non-parametric regression method, with priors imposed directly on the covariance function of the resulting approximation (Neal, 1996, 1997, 1999). Given the training inputs $x^{(1)}, \ldots, x^{(n)}$ and the new input $x^{(n+1)}$, a covariance function can be used to compute the $n+1$ by $n+1$ covariance matrix of the associated targets $y^{(1)}, \ldots, y^{(n)}, y^{(n+1)}$. The predictive distribution for $y^{(n+1)}$ is obtained by conditioning on the known targets, giving a Gaussian distribution with the mean and the

variance given by

$$E_y[y|x^{(n+1)}, \theta, D] = k^{\mathrm{T}} C^{-1} y^{(1,\ldots,n)} \tag{6}$$

$$\mathrm{Var}_y[y|x^{(n+1)}, \theta, D] = V - k^{\mathrm{T}} C^{-1} k, \tag{7}$$

where C is the n by n covariance matrix of the observed targets, $y^{(1,\ldots,n)}$ is the vector of known values for these targets, k is the vector of covariances between $y^{(n+1)}$ and the known n targets, and V is the prior variance of $y^{(n+1)}$. We used a simple quadratic covariance function producing smooth functions and hierarchical prior for parameters related to the effects of variables, with elements

$$C_{ij} = \eta^2 \exp\left(-\sum_{u=1}^{p} \rho_u^2 (x_u^{(i)} - x_u^{(j)})^2 \right) + \delta_{ij} J^2 + \delta_{ij} \sigma_e^2. \tag{8}$$

The first term of this covariance function expresses that the cases with nearby inputs should have highly correlated outputs. The η parameter gives the overall scale of the local correlations. The ρ_u parameters are multiplied by the coordinate-wise distances in the input space and thus allow for different distance measures for each input dimension. The second term is the jitter term, where δ_{ij} is the Kronecker delta. It is used to improve matrix computations by adding a constant term to the residual model. The third term is the residual model. The residual was modeled with Student's t-distribution expressed as a mixture of normal distributions with unknown width and degrees of freedom, and recipe dependent parameters. Posterior and predictive distributions were computed with MCMC methods. See References (Neal, 1999; Vehtari and Lampinen, 2001) for details of the algorithms, models, priors and MCMC parameters. To assess usefulness of the image analysis features we constructed a model with all the features and another model with all the features except the image analysis features.

5 Model comparison

To test whether the model with the image analysis features gives better predictions than the model without them, we need to make model comparison comparing the predictive performance of the models.

In prediction and decision problems, it is natural to assess the predictive ability of the model by estimating the expected utilities, that is, the relative values of consequences of using the model (Bernardo and Smith, 1994). The goal is to estimate how good predictions the model can make if used for future measurements. Expected utilities can be estimated using cross-validation (CV) predictive densities (Vehtari and Lampinen, 2002, 2003). Cross-validation approximates the external validation, that is, testing the model with real future measurements. To simulate the fact that the future observations are not in the training data, the ith observation $(x^{(i)}, y^{(i)})$ in the training data is left out, and then the predictive distribution for $y^{(i)}$ is computed with a model that is fitted to all of the observations except $(x^{(i)}, y^{(i)})$.

To reduce the computational burden we used k-fold-CV, in which data is divided in k parts, and each part is left out in turn.

We used predictive likelihood utility for the next observation

$$u_M = p(y^{(n+1)}|x^{(n+1)}, D, M) \tag{9}$$

which measures how well the model estimates the whole predictive distribution and not just some summary statistic of the prediction, and thus is especially useful in model comparison. We prefer to report the expectation of the factor of predictive likelihoods

$$\bar{u}_{M_1,M_2} = \mathrm{E}[u_{M_1}/u_{M_2}]. \tag{10}$$

This value is not directly dependent on the size of the training data n and common scale term in both u_{M_1} and u_{M_2} cancels out. An often used alternative is to report the expected predictive deviances $\bar{D}_{M_i} = -2n\,\mathrm{E}[u_{M_i}]$ (Spiegelhalter et al., 2002, e.g.).

To describe the associated uncertainty in the estimates, we used a quick and generic approach based on the Bayesian bootstrap for obtaining samples from the distributions of the expected utility estimates (Vehtari and Lampinen, 2002). These samples can be used to compute, for example, 95% credible intervals for the estimates or the probability of one model having a better expected predictive likelihood than the alternative model.

Furthermore, cross-validation predictive densities were used to compute expected predictive accuracy of the models in units of quality measurements (e.g., in MPa for compressive strength), and thus the concrete expert could assess the practical usefulness of the models.

6 Results

Figure 4 shows the distributions of the estimates of the factors of the predictive likelihoods (Eq. (10)). If models had same predictive performance, this factor would be 1. Factor values larger than 1 imply that the model with the image analysis features gives better predictions. The distributions of the estimates describe the associated uncertainties in these model comparisons. Figure 4 also shows the probabilities (p) that the model with image analysis features makes better predictions than the model without these features. Based on these, it is statistically likely that image analysis features are useful for predicting air percentage and compressive strength. Furthermore, it is statistically very likely that these features are useful for predicting flow value and bleeding. More detailed analysis of the effects of the different features and comparison to expert knowledge can be found in references (Järvenpää, 2001; Vehtari, 2001).

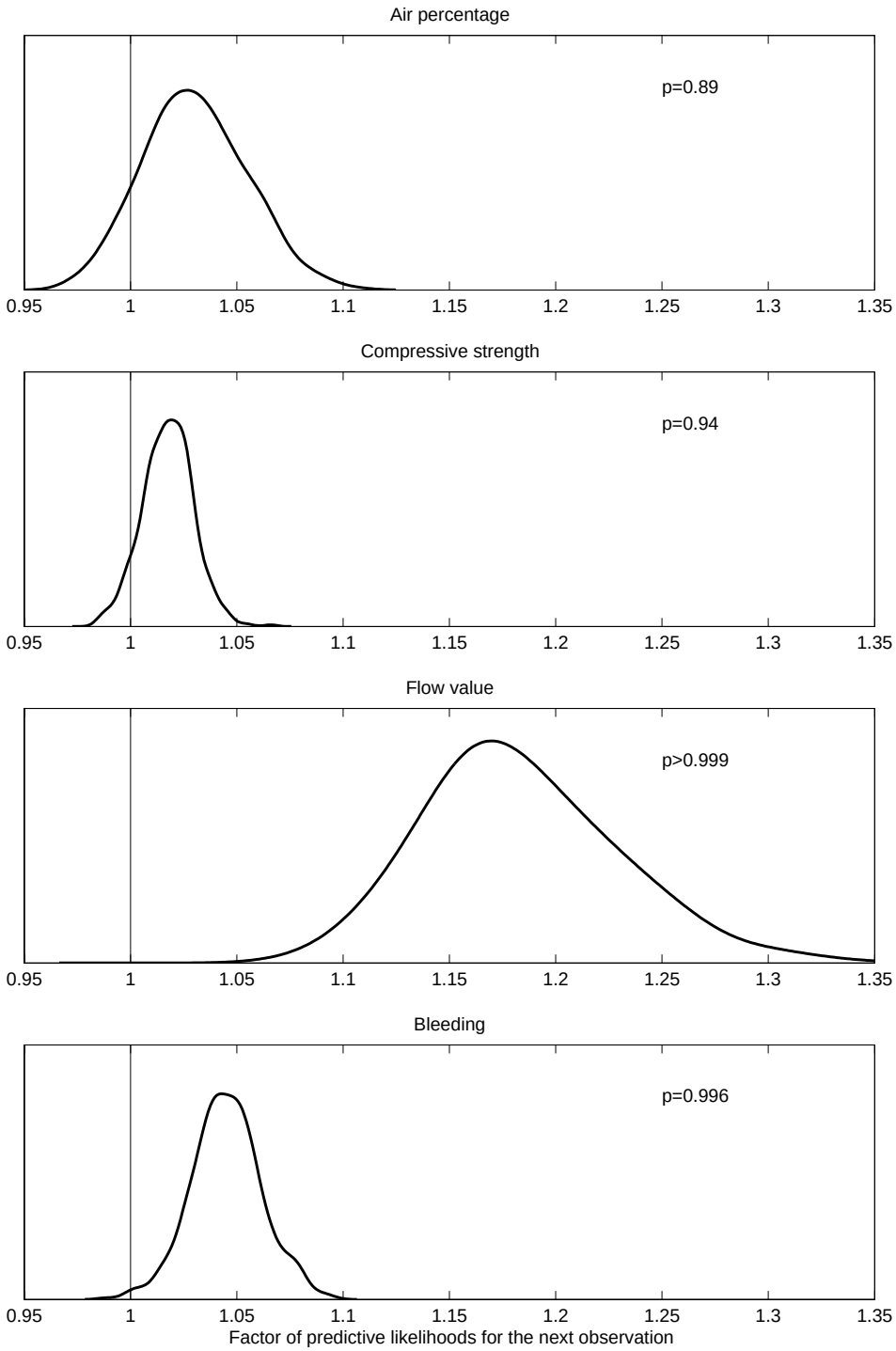


Figure 4: Distributions of estimates of factors of predictive likelihoods and the probability that the model with image analysis features gives better predictions.

7 Conclusion

We have presented models for predicting the properties of concrete which use visual features in addition to physical measurements as input variables. The visual features were found to be useful in improving the predictions of the models. By using the models and conclusions based on them it was possible to reduce the proportion of the natural gravel from 50% to 5%-20% and achieve 5%-15% savings in concrete manufacturing (Järvenpää, 2003).

Acknowledgments

This study was partly funded by TEKES Grant 40888/97 (Project *PROMISE, Applications of Probabilistic Modeling and Search*), Graduate School in Electronics, Telecommunications and Automation (GETA), Graduate School in Computational Methods of Information Technology (ComMIT) and the Nokia Foundation. The authors would like to thank Dr. H. Järvenpää for providing her expertise in this concrete problem.

References

Bernardo JM, Smith AFM (1994). Bayesian Theory. John Wiley & Sons.

Järvenpää H (2001). Quality characteristics of fine aggregates and controlling their effects on concrete. Acta Polytechnica Scandinavica, Civil Engineering and Building Construction Series No. 122. The Finnish Academy of Technology.

Järvenpää H (2003). Product Director, Lohja Rudus. Personal communication.

Masad E, Button JW (2000). Unified imaging approach for measuring aggregate angularity and texture. Computer Aided Civil and Infrastructure Engineering, 15(4):273–280.

Neal RM (1996). Bayesian Learning for Neural Networks. Springer.

Neal RM (1997). Monte Carlo implementation of Gaussian process models for Bayesian regression and classification. Technical Report 9702, Dept. of Statistics, University of Toronto.

Neal RM (1999). Regression and classification using Gaussian process priors (with discussion). In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds., Bayesian Statistics 6. Oxford University Press, pp. 475–501.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002). Bayesian measures of model complexity and fit (with discussion). Journal of the Royal Statistical Society Series B Statistical Methodology, 64(3):583–639.

Tuceryan M, Jain A (1993). Texture analysis. In: Chen C, Pau L, Wang P, eds., Handbook of Pattern Recognition and Vision. World Scientific, pp. 235–276.

Vehtari A (2001). Bayesian Model Assessment and Selection Using Expected Utilities. Dissertation for the degree of Doctor of Science in Technology, Helsinki University of Technology.

Vehtari A, Lampinen J (2001). On Bayesian model assessment and choice using cross-validation predictive densities. Technical Report B23, Helsinki University of Technology, Laboratory of Computational Engineering.

Vehtari A, Lampinen J (2002). Bayesian model assessment and comparison using cross-validation predictive densities. Neural Computation, 14(10):2439–2468.

Vehtari A, Lampinen J (2003). Expected utility estimation via cross-validation. In: Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, Smith AFM, West M, eds., Bayesian Statistics 7. Oxford University Press, pp. 701–710.