Model formulations for Gaussian animal model

A Gaussian animal model can be formulated in two alternative ways, both fitting the INLA framework.

**Model formulation 1 (MF1):** Likelihood \( y_i | \eta_i \sim N(\eta_i, \sigma^2_e) \) and latent field \( \eta_i = \beta_0 + z_i^T \beta + u_i + \epsilon_i \), where the variance of \( \epsilon \) is fixed to a small value.

**Model formulation 2 (MF2):** Likelihood \( y_i | \eta_i \sim N(\eta_i, \sigma^2_{small}) \), i.e. the variance of the likelihood is fixed to a small value, and latent field \( \eta_i = \beta_0 + z_i^T \beta + u_i + \epsilon_i \), where the variance of \( \epsilon \) is \( \sigma^2_e \). The \( \sigma^2_{small} \) can be interpreted as measurement uncertainty.

When estimating the narrow sense heritability, \( h^2 \), in the Gaussian case, we use model formulation MF2, which out of convenience is parametrized with \( (\sigma^2_u, h^2) \) instead of \( (\sigma^2_u, \sigma^2_e) \). Further, \( (\sigma^2_u, h^2) \) is given a prior such that it corresponds to the prior of \( (\sigma^2_e, \sigma^2_u) \), hence, the same prior under two different parametrizations.

The DIC is based on evaluating the likelihood, and is not invariant with respect to parametrization (Spiegelhalter et al. 2002). Using model formulation MF2, i.e. a fixed small variance for the likelihood does not work numerically; almost all models get the same DIC to the precision given by INLA. So if DIC needs to be calculated the animal model has to be formulated in an alternative way (in the INLA framework), where the variance of \( \epsilon \) is fixed to a small value in the latent field, i.e. using MF1. Both model formulations coincide if the same priors are used for the hyper-parameters \( (\beta, \sigma^2_e, \sigma^2_u) \), and are latent Gaussian fields with only two non-Gaussian parameters, namely \( \theta = (\sigma^2_u, \sigma^2_e) \). For MF1 \( \epsilon \) can be omitted from the model. It is included here to be consistent with MF2. Both model formulations have their numerical advantages depending on the aim of the analysis. However, we have to be cautious which model formulation we use depending on the purpose of the analysis.

To summarize, when \( u_i, \sum_{i \in C} w_i u_i, \beta \) or \( \sigma^2_u \) is of interest both MF1 and MF2 might be used. If \( \sigma^2_e \) or DIC is the aim of the analysis MF1 has to be used, while MF2 with parametrization \( (\sigma^2_e, h^2) \) has to be used if \( h^2 \) is of interest. Hence we might have to fit two (INLA) models to get all estimates of interest.
Literature Cited