White, black, or rather grey? Choosing color for the predictive model for human Puumala orthohantavirus infections driven by a common forest rodent

Orestis Kazasidis and Jens Jacob

Julius Kühn Institute (JKI) – Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Horticulture and Forests, Vertebrate Research, Münster

E-mail of corresponding author: orestis.kazasidis@julius-kuehn.de

The Puumala orthohantavirus (PUUV) is the most common hantavirus in Europe. It is transmitted to humans by infected bank voles, a common forest rodent that can cause significant damage to young trees. PUUV infection in humans can cause severe symptoms, such as kidney failure. Forestry works are particularly at risk. Monitoring the distribution of the virus both in the bank vole populations and in humans shows that it is present in several states of western and southern Germany. According to data from the Robert Koch Institute from 2010 to 2019, there have been >1000 reported human PUUV infections per year in 2010, 2012, 2017, and 2019, with a yearly incidence of 2.29, 2.94, 1.37, and 1.25, respectively. In 2021, there are already 813 reported infections (incidence 0.98, status 05-07-2021).

Weather parameters, such as temperature and precipitation, and the proportion of deciduous forest have been identified as strong predictors for human PUUV infections. The weather parameters are linked to the beech mast intensity of the following year, which determines the food availability and governs the growth of the bank vole populations. In turn, the deciduous forest is the natural habitat of the bank voles, and its proportion is a proxy for the probability of human-bank vole interaction.

Previous work in our group led to two distinct predictive models for human PUUV infections, for districts in northern and southern Germany, respectively, both based on Classification and Regression Trees (CART). These models are “white boxes”, meaning that their function can be directly inspected and analyzed. In autumn 2020, we started a project to improve these models.

During this talk, we will explore the predictors, by observing characteristic 2D and 3D plots. We will use these plots to interpret the results of white-box approaches. Next, we will discuss the potential of using “black-box” techniques, such as the random forest and the artificial neural networks, and present countermeasures to gain intuition in their internal operation; rendering them rather “grey”. The results can be used to improve prediction of outbreaks of disease and forest pest rodent dynamics alike. This is highly important to support strategies and decisions for the protection of plants and human health.

This project was commissioned by the German Environment Agency and funded by the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety within the departmental research plan - research code 3720 48 401 0.