Regression models for Cylindrical data in Psychology

Abstract

Cylindrical data are multivariate data which consist of a directional, in this paper circular, and a linear component. Examples of cylindrical data in psychology include human navigation (direction and distance of movement), eye-tracking research (direction and length of saccades) and data from an interpersonal circumplex (location and intensity on the IPC). In this paper we adapt four models for cylindrical data to include a regression of the circular and linear component onto a set of covariates. Subsequently, we illustrate how to fit these models and interpret their results on a dataset on the interpersonal behavior of teachers.

Keywords: cylindrical data, regression, interpersonal behavior

Regression models for Cylindrical data in Psychology

In the social sciences the use of cylindrical data is very common. Such data consist of a linear and a circular component. Gurtman (2011) refers to such data as vectors, with a directional measure (i.e., the circular component) and a measure indicating the magnitude (i.e., the linear component). Many established models in psychology are often referred to as circular or circumplex models, but those models are cylindrical. Examples of such cylindrical models are the interpersonal circle/circumplex (Leary, 1957; Wiggins, 1996; Wubbels, Brekelmans, Den Brok, & Van Tartwijk, 2006), the circumplex of affect (Russell, 1980), the circumplex of human emotion (Plutchik, 1997) or the model of human values (Schwartz, 1992).

Also, many of the more recent types of data that are studied in psychology are cylindrical. For example, research on human navigation uses data where distance (i.e., the linear component) and direction (i.e., the circular component) are of interest (Chrastil & Warren, 2017) or in eye-tracking, the saccade data also consist of both the direction (i.e., the circular variable) and the duration (i.e., the linear variable) (e.g., Rayner (2009)). Apart from the social sciences, data with a circular and linear component more commonly occur in meteorology (García-Portugués, Crujeiras, & González-Manteiga, 2013), ecology (García-Portugués, Barros, Crujeiras, González-Manteiga, & Pereira, 2014) or marine research (Lagona, Picone, Maruotti, & Cosoli, 2015)

Up until now researchers studying cylindrical data had to rely on linear statistical methods to analyze their research results. However, lately more and more of these researchers acknowledge that linear methods are not sufficient and call for new methods (Gurtman, 2011; Pennings, 2017b; Wright, Pincus, Conroy, & Hilsenroth, 2009) that take into account both the circular and the linear component of these data.

The aim in the present paper is twofold. Firstly, we intend to fill the above mentioned gap in the literature by showing that the use of cylindrical models can benefit the analysis of circumplex data and cylindrical data in psychology in general. More specifically we will show these benefits for interpersonal teacher data from the field of educational psychology. Apart from modelling the dependence between the linear and circular component of a cylindrical variable we would also like to predict the two components from a set of covariates in a regression model. Our second aim therefore is to adapt several existing cylindrical models in such a way that they include a regression of both the linear and circular component of a cylindrical variable onto a set of covariates. These adapted cylindrical models are then used to analyse the teacher data.

Modelling Framework

Data that consist of a linear variable and a circular variable are called cylindrical data. The circular variable is different from the linear variable in the sense that it is measured on a different scale. Figure 1 shows the difference between a circular scale (right) and a linear scale (left). The most important difference is that on a circular scale the datapoints $0°$ and $360°$ are connected and in fact represent the same number while on a linear scale the two ends, $-\infty$ and ∞, are not connected and consequently the values $0°$ and $360°$ are located on different places on the scale. Both circular data and cylindrical data require special analysis methods due to this periodicity in the scale of a circular variable (see e.g. Fisher (1995) for an introduction to circular data and Mardia and Jupp (2000), Jammalamadaka and Sengupta (2001) and Ley and Verdebout (2017) for a more elaborate overview).

[Figure 1 about here]

From now on we will refer to the circular random variable as Θ, the circular component, and the linear random variable as Y, the linear component, with realizations θ_i and y_i for each measurement $i = 1, \dots, n$ where n is the sample size. The circular component is measured in radians and consequently takes values in the interval $[0, 2\pi)$. Note that the round brackets mean that 2π is not included in the interval since this represents the same value as 0 as a result of periodicity. Radians can be converted to degrees using the relation 1 rad $= 1 * \frac{180°}{\pi}$. A circular variable is usually described using its first two moments, circular location or mean μ and mean resultant length ρ[1]. The sample values of these two parameters are referred to as $\bar{\theta}$ and $\hat{\rho}$. The mean resultant length lies between 0, meaning the data is not concentrated at all *i.e.* spread over the entire circle, and 1, meaning all data is concentrated at a single point on the circle. The linear component is measured on the real line and can take values on $(-\infty, \infty)$ or we can constrain it to the positive real line $(0, \infty)$. Depending on which distributional assumptions are used we describe the distribution of the linear component using a mean μ and variance σ^2 or a scale ν and shape α parameter. When skewed distributions are used for either the circular or linear component we may also include a skewness parameter λ.

In a cylindrical framework we model the location μ_c of the circular variable and the mean μ_l or scale ν of the linear variable. Note that to prevent confusion we distinguish between the circular location and linear mean using subscripts. In this paper we use covariates to predict the circular and linear component using the following general type of prediction equations for a cylindrical model with q covariates for the linear component and p covariates for the circular component:

$$\hat{\theta}_i = g(\beta_0 + \beta_1 z_1 + \cdots + \beta_q z_p), \tag{1}$$

$$\hat{y}_i = h(\gamma_0 + \gamma_1 x_1 + \cdots + \gamma_q x_q), \tag{2}$$

where $g()$ and $h()$ are link functions, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)$ is a vector of regression coefficients for the circular variable and $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_q)$ is a vector of regression coefficients for the linear variable. The link functions that are chosen depend on the type of distribution we choose for the linear and circular component. In this paper we use either an identity link or an exponential function (when modelling a scale parameter) for the linear component. For the

[1] In the von Mises distribution, a common distribution for a circular outcome, we have a concentration parameter κ that is related to ρ as $\rho = A_1(\kappa)$, where $A_1\kappa = I_1(\kappa)/I_0(\kappa)$ and $I_0()$ and $I_1()$ are modified Bessel functions of order 0 and 1, respectively.

circular component the link functions are specific to circular data and include arctangent functions. The specific type of link function and prediction equations used will be introduced in the respective descriptions of the models used in this paper.

In addition to modelling the circular and linear component of a cylindrical variable separately using covariates we also model the relation between them. In the literature, several methods have been put forward to do so. Some of these are based on regressing the linear component onto the circular component by including $\sin(\theta_i)$ and $\cos(\theta_i)$ into the prediction equation in (2) (Johnson & Wehrly, 1978; Mardia & Sutton, 1978; Mastrantonio, Maruotti, & Jona-Lasinio, 2015). The linear component is thus modeled using the sine and cosine of the circular component. Others model the relation in a different way, e.g. by specifying a multivariate model for the linear and circular variable and modelling their covariance matrix (Mastrantonio, 2018) or by proposing a joint cylindrical distribution. For example, Abe and Ley (2017) introduce a cylindrical distribution based on a Weibull distribution for the linear component and a sine-skewed von Mises distribution for the circular component and link these through their respective shape and concentration parameters.

Four Cylindrical Regression Models

One of the goals of this paper is to show the benefits of cylindrical methods for the analysis of circumplex data and cylindrical data in psychology in general. To do so we focus on four cylindrical models. The models were selected for their relatively low complexity and the ease with which a regression structure could be incorporated. But also because they show different ways of modelling the linear and circular component and thereby illustrate a wider range of cylindrical models available in the literature. As outlined in the previous section, the cylindrical models contain a set of q predictors $\boldsymbol{x} = x_1, \ldots, x_q$ and p predictors $\boldsymbol{z} = z_1, \ldots, z_p$ for the linear and circular components, Y and Θ, respectively. The first two models are based on a construction by Mastrantonio et al. (2015), while the other models are extensions of the models from Abe and Ley (2017) and Mastrantonio (2018). The four cylindrical models are introduced separately in the subsections below. However, to provide a more succinct overview and comparison of the four models, Table 1 gives an overview of the similarities and differences between the models.

[Table 1 about here]

The Modified Circular-Linear Projected Normal (CL-PN) and Modified Circular-Linear General Projected Normal (CL-GPN) Models. Following Mastrantonio et al. (2015) we consider two models where the prediction equation for the linear component is specified by combining (2), with an identity link function and including $\sin(\theta_i)$ and $\cos(\theta_i)$ as follows:

$$\hat{y}_i = \gamma_0 + \gamma_{cos} * \cos(\theta_i) * r_i + \gamma_{sin} * \sin(\theta_i) * r_i + \gamma_1 * x_1 + \cdots + \gamma_q * x_q, \tag{3}$$

where r_i is a realization of the unobserved random variable $R \geq 0$ that will be introduced below, $\gamma_0, \gamma_{cos}, \gamma_{sin}, \gamma_1, \ldots, \gamma_q$ are the intercept and regression coefficients and $x_1, \ldots, x_q$ are the q covariates for the prediction of the linear component. We assume a normal $N(\mu_l, \sigma^2)$ distribution for the linear component.

For the circular component we assume either a projected normal (PN) or a general projected normal (GPN) distribution. These distributions arise from a projection of a distribution defined in bivariate space onto the circle. Figure 2 represents this projection. In the left plot of Figure 2 we see realizations (s_i^I, s_i^{II}) from the bivariate normal variable $\boldsymbol{S}$ that in the middle plot are projected to form the circular component Θ in the right plot. Mathematically the relation between $\boldsymbol{S}$ and Θ is defined as follows:

$$\boldsymbol{S} = \begin{bmatrix} S^I \\ S^{II} \end{bmatrix} = \begin{bmatrix} R\cos(\Theta) \\ R\sin(\Theta) \end{bmatrix}, \tag{4}$$

where $R = \mid\mid \boldsymbol{S} \mid\mid$, the Euclidean norm of $\boldsymbol{S}$, that is represented by the lines connecting the bivariate datapoints to the origin in the middle plot. We call $\boldsymbol{S}$ the augmented representation of the circular component. It is a variable that in contrast to Θ is not observed and thus considered latent or auxiliary. This then means that we do not model Θ directly but indirectly through $\boldsymbol{S}$.

[Figure 2 about here]

For both the PN and GPN distributions the circular location parameter $\mu_c \in [0, 2\pi)$ is modeled.[2] The prediction equation for the circular component is specified by using a double arctangent link function in (1) as follows:

$$\hat{\theta}_i = \text{atan2}(\boldsymbol{\beta^{II}} \boldsymbol{z}_i, \boldsymbol{\beta^{I}} \boldsymbol{z}_i) \tag{5}$$

where $\boldsymbol{\beta}^I = (\beta_0^I, \beta_1^I, \ldots, \beta_p^I)$ and $\beta^{II} = (\beta_0^{II}, \beta_1^{II}, \ldots, \beta_p^{II})$ are vectors with intercepts and regression coefficients for the prediction of $S^{\boldsymbol{I}}$ and S^{II} and $\boldsymbol{z}_i$ is a vector with predictor values for each individual $i \in 1, \ldots, n$ where n is the sample size. Note that as a result of the augmented representation of the circular component we have two sets of regression coefficients and intercepts, in contrast to a single set in (1). This leads to problems when we want to interpret the effect of a covariate on the circle. A circular regression line for simulated data is shown in Figure 3, with covariate values on the x-axis and the predicted circular component on the y-axis. As can be seen it is of a non-linear character meaning that the effect of a covariate is different at different values of the covariate. A circular regression line is usually described by the slope at the inflection point, the point at which the slope of the regression line starts flattening off (indicated with a square in Figure 3). By default, the parameters from the PN and GPN models do not directly describe this inflection point. For the PN distribution however, Cremers et al. (2018b) solved this interpretation problem. They introduce a new parameter b_c that describes the slope at the inflection point of the regression line. For the GPN distribution the interpretation problem however remains.

[2] Note that for the CL-GPN model the circular location parameter also depends on the variance-covariance matrix and the circular predicted values should be computed using numerical integration or Monte Carlo methods because a closed form expression for the mean direction is not available.

[Figure 3 about here]

The main difference between the PN and GPN distribution lies in the definition of their covariance matrix. For the PN distribution this is an identity matrix, causing the distribution to be unimodal and symmetric, whereas for the GPN distribution $\mathbf{\Sigma} = \begin{bmatrix} \tau^2 + \xi^2 & \xi \\ \xi & 1 \end{bmatrix}$ where $\xi, \tau \in (-\infty, +\infty)$, allowing for multimodality and asymmetry/skewness as illustrated in Figure 4.

[Figure 4 about here]

Both the CL-PN and CL-GPN models are estimated using Markov Chain Monte Carlo (MCMC) methods based on Nuñez-Antonio, Gutiérrez-Peña, and Escarela (2011), Wang and Gelfand (2013) and Hernandez-Stumpfhauser, Breidt, and Van der Woerd (2016) for the regression of the circular component. A detailed description of the Bayesian estimation and MCMC samplers can be found in the Supplementary Material.

The Modified Abe-Ley Model. This model is an extension of the cylindrical model introduced in Abe and Ley (2017) to the regression context. The circular dependence between the linear and circular component, Y and Θ, is defined through a joint density function as follows:

$$f(\theta, y) = \frac{\alpha\nu^\alpha}{2\pi\cosh(\kappa)}(1 + \lambda\sin(\theta - \mu_c))y^{\alpha-1}\exp[-(\nu y)^\alpha(1 - \tanh(\kappa)\cos(\theta - \mu_c))], \quad (6)$$

where $\alpha > 0$ is a linear shape parameter, $\kappa > 0$ and $\lambda \in [-1, 1]$ are circular concentration and skewness parameters respectively. The Abe-Ley density thus concerns a combination of a Weibull distribution, with scale parameter $\nu > 0$ and shape parameter α, for the linear component and a sine-skewed von Mises distribution, with location parameter $\mu_c \in [0, 2\pi)$, concentration parameter $\kappa > 0$ and skewness $\lambda \in [-1, 1]$, for the circular component. In contrast to the CL-PN and CL-GPN models, the linear component Y is in this model defined only on the positive real half-line $[0, +\infty)$ and thus can not be negative. Our modification of the Abe-Ley model occurs at the level of the linear scale parameter $\nu > 0$ and circular location parameter $\mu_c \in [0, 2\pi)$, both of which we will model using covariates.

The prediction equation for the circular component equals:

$$\hat{\theta}_i = \beta_0 + 2\tan^{-1}(\boldsymbol{\beta}\boldsymbol{z}_i) \tag{7}$$

Note that we have taken out the intercept β_0 from the link function $g() = 2\tan^{-1}$ in (1). The vector $\boldsymbol{\beta}$ thus does not contain an intercept. We do not directly predict the linear component. The conditional distribution for the linear component is Weibull, meaning that we can use methods from survival analsis to interpret the effect of a predictor. In survival analysis a "survival" function is used in which time is plotted against the probability of survival of subjects suffering from a specific medical condition. This probability is computed using the "survival-function" defined as

$$\exp(-\alpha y_i^{\hat{\nu}_i(1-\tanh(\kappa)\cos(\theta_i-\hat{\theta}_i))^{1/\alpha}}), \tag{8}$$

with $\hat{\nu}_i = \exp(\boldsymbol{\gamma}\boldsymbol{x}_i)$. Note that $\boldsymbol{x}_i$ also includes a 1 to be able to estimate the intercept γ_0. From the survival function we also see that the circular concentration parameter κ and the linear shape parameter α regulate the circular-linear dependence in the Abe-Ley model. In the exponent or power of the linear outcome, the prediction error on the circle, $\cos(\theta_i - \hat{\theta}_i)$, is multiplied by $\tanh \kappa$, and raised to the power $1 - \alpha$.

We can use numerical optimization (Nelder-Mead) to find solutions for the maximum likelihood (ML) estimates for the parameters of the model.

Modified Joint Projected and Skew Normal Model (GPN-SSN). This model is an extension of the cylindrical model introduced by Mastrantonio (2018) to the regression context. Although the model may contain several circular and linear components we will restrict ourselves to one circular and one linear component in this paper. The circular component is modelled by a GPN distribution, as in the CL-GPN model, while the linear component is modelled by a skew normal distribution (Sahu, Dey, & Branco, 2003). Because the GPN distribution is modelled using a so-called augmented representation (see (4)) it is convenient to use a similar tactic for modelling the skew normal distribution. As in Mastrantonio (2018), dependence between the linear and circular component is created by modelling the augmented representation of Θ and Y together in a 3 dimensional normal distribution. The joint density of the model is then represented by:

$$f(\theta, r, y, d) = 2\phi_3((\boldsymbol{s}^t, y^t)^t \mid \boldsymbol{M} + (0, 0, \lambda d)^t, \boldsymbol{\Sigma})\phi(d \mid 0, 1)r, \tag{9}$$

where ϕ represents the normal probability density function, $\boldsymbol{s} = (r(\cos(\theta), \sin(\theta)))^t$, d is the augmented representation of the linear component, $\boldsymbol{M} = \boldsymbol{B}^t\boldsymbol{X}$, $\boldsymbol{B}_{(g+1)\times 3}$ is defined below, $\boldsymbol{X}$ is a design matrix and $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_s & \boldsymbol{\Sigma}_{sy} \\ \boldsymbol{\Sigma}_{sy}^t & \boldsymbol{\Sigma}_y \end{pmatrix}$ a variance-covariance matrix. The matrices $\boldsymbol{\Sigma}_s$ and $\boldsymbol{\Sigma}_y$ are the covariance matrices for the variances of and covariances between the augmented representations of the circular and linear component respectively. Note that because we only have one circular and one linear component in this paper, $\boldsymbol{\Sigma}_s$ is a two by two matrix and $\boldsymbol{\Sigma}_y$ is a scalar representing the variance of the augmented linear component. The matrix $\boldsymbol{\Sigma}_{sy}$ contains covariances between the augmented representations of the circular and linear component. The matrix with regression coefficients and intercepts, $\boldsymbol{B}$ is defined as follows:

$$\boldsymbol{B} = \begin{bmatrix} \beta_{0_{sI}} & \beta_{0_{sII}} & \beta_{0_y} \\ \beta_{1_{sI}} & \beta_{1_{sII}} & \beta_{1_y} \\ \vdots & \vdots & \vdots \\ \beta_{g_{sI}} & \beta_{g_{sII}} & \beta_{g_y} \end{bmatrix}. \tag{10}$$

Note that in contrast to the previous three models where we use p and q covariates for the circular and linear component respectively, we use $g = p = q$ covariates for both components in the GPN-SSN model. Unlike the other models, where μ_l and μ_c are modelled separately, the GPN-SSN thus has a shared mean vector $\boldsymbol{M}$ and variance-covariance matrix $\boldsymbol{\Sigma}$ for the linear and circular component, much like having multiple outcomes in a MANOVA (multivariate analysis of variance) model. This implies that the covariances in the variance-covariance matrix describe the dependence between the circular and linear component.

The prediction equation for the circular component[3] is similar to (5) and equal to:

$$\hat{\theta}_i = \text{atan2}(\boldsymbol{\beta^{II}} \boldsymbol{z}_i, \boldsymbol{\beta^{I}} \boldsymbol{z}_i), \tag{11}$$

where $\boldsymbol{\beta^I} = (\beta_{0_{sI}}, \beta_{1_{sI}}, \ldots, \beta_{g_{sI}})$ and $\boldsymbol{\beta^{II}} = (\beta_{0_{sII}}, \beta_{1_{sII}}, \ldots, \beta_{g_{sII}})$. For the linear component we subsititute $h()$ in (2) for an identity link and $\boldsymbol{\gamma}$ for $\boldsymbol{\beta}$:

$$\hat{y}_i = \boldsymbol{\beta} \boldsymbol{x}_i, \tag{12}$$

where $\boldsymbol{x}_i = \boldsymbol{z}_i$ and $\boldsymbol{\beta} = (\beta_{0_y}, \beta_{1_y}, \ldots, \beta_{g_y})$.

We estimate the model using MCMC methods. A detailed description of these methods is given in the Supplementary Material.

Model Fit Criterion

For the four cylindrical models we focus on their out-of-sample predictive performance to determine the fit of the model. To prevent possible problems concerning overestimation, we use k-fold cross-validation and split our data into 10 folds. Each of these folds (10 % of the sample) is used once as a holdout set and 9 times as part of a training set. The analysis will thus be performed 10 times, each time on a different training set.

A proper criterion to compare out-of-sample predictive performance is the Predictive Log Scoring Loss (PLSL) (Gneiting & Raftery, 2007). The lower the value of this criterion, the better the predictive performance of the model. Because the joint density and thus also the likelihood for the modified GPN-SSN model is not available in closed form (Mastrantonio, 2018) we compute the PLSL for the circular and linear component separately for all models. Using ML estimates this criterion can be computed as follows for the circular

[3] Note that for the GPN-SSN model the predicted circular component also depends on the variance-covariance matrix and the circular predicted values should be computed using numerical integration or Monte Carlo methods because a closed form expression for the mean direction is not available.

and linear component:

$$PLSL_c = -2\sum_{i=1}^{M} \log l(\theta_i \mid \hat{\boldsymbol{\vartheta}}),$$

$$PLSL_l = -2\sum_{i=1}^{M} \log l(y_i \mid \hat{\boldsymbol{\vartheta}}),$$

where l is the model likelihood, M is the sample size of the holdout set, y_i and θ_i are the i^{th} datapoints from the holdout set and $\hat{\boldsymbol{\vartheta}}$ are the ML estimates of the model parameters. Using posterior samples the criterion is similar to the log pointwise predictive density (lppd) (Gelman et al., 2014, p. 169) and can be computed for the circular and linear component as:

$$PLSL_c = -2\frac{1}{B}\sum_{j=1}^{B}\sum_{i=1}^{M} \log l(\theta_i \mid \boldsymbol{\vartheta}^{(j)}),$$

$$PLSL_l = -2\frac{1}{B}\sum_{j=1}^{B}\sum_{i=1}^{M} \log l(y_i \mid \boldsymbol{\vartheta}^{(j)}),$$

where B is the amount of posterior samples and $\boldsymbol{\vartheta}^{(j)}$ are the posterior estimates of the model parameters for the j^{th} iteration. Note that although we fit the CL-PN, CL-GPN and GPN-SSN models using Bayesian statistics, we do not take prior information into account when assessing model fit with the PLSL. According to Gelman et al. (2014) this is not necessary since we are assessing the fit of a model to data, the holdout set, only. They argue that the prior in such case is only of interest for estimating the parameters of the model but not for determining the predictive accuracy.

For each of the four cylindrical models and for each of the 10 cross-validation analyses we can then compute a PLSL for the circular and linear component by using the conditional log-likelihoods of the respective component (see Supplementary Material for a definition of the loglikelihoods). To evaluate the predictive performance we average across the PLSL criteria of the cross-validation analyses. We also assess the cross-validation variability by means of the standard deviations of the PLSL criteria.

Teacher Data

The motivating example for this article comes from the field of educational psychology and was collected for the studies on classroom climate of Van der Want (2015), Claessens (2016) and Pennings (2017a). An indicator of the quality of the classroom climate is the students' perception of their teachers' interpersonal behavior. These interpersonal perceptions, both in educational psychology as well as in other areas of psychology, can be measured using circumplex measurement instruments (see Horowitz and Strack (2011) for an overview of many such instruments).

The circumplex data used in this paper are measured using the Questionnaire on Teacher Interaction (QTI) (Wubbels et al., 2006) which is one such circumplex measurement instrument. The QTI is designed to measure student perceptions of their teachers' interpersonal behavior and contains items that load on two interpersonal dimensions: Agency and Communion. Agency refers to the degree of power or control a teacher exerts in interaction with his/her students. Communion refers to the degree of friendliness or affiliation a teacher conveys in interaction with his/her students. The loadings on the two dimensions of the QTI can be placed in a two-dimensional space formed by Agency (vertical) and Communion (horizontal), see Figure 5. This space is called the interpersonal circle/circumplex (IPC) and different parts of this space are characterized by different teacher behavior, e.g. "helpful" or "uncertain". The IPC is "a continuous order with no beginning or end" (Gurtman, 2009, p. 2). We call such ordering a circumplex ordering and the IPC is therefore often called the interpersonal circumplex. The ordering also implies that scores on the IPC could be viewed as a circular variable. This circular variable represents location on the IPC of the interpersonal behavior that a teacher shows towards his/her students.

[Figure 5 about here]

Cremers et al. (2018a) explain the circular nature of the IPC data and analyze them as such using a circular regression model. The two dimension scores, Agency and Communion, can be converted to a circular score using the two-argument arctangent function in (13), where A represents a score on the Agency dimension and C represents a score on the Communion dimension[4]. Note that when placing a unit circle on Figure 5 we see that the Agency dimension is related to the sine of the circular score and the Communion dimension is related to the cosine of the circular score.

$$\theta = \text{atan2}\left(A,\ C\right) = \begin{cases} \arctan\left(\frac{A}{C}\right) & \text{if} \quad C > 0 \\ \arctan\left(\frac{A}{C}\right) + \pi & \text{if} \quad C < 0 \ \& \ A \geq 0 \\ \arctan\left(\frac{A}{C}\right) - \pi & \text{if} \quad C < 0 \ \& \ A < 0 \\ +\frac{\pi}{2} & \text{if} \quad C = 0 \ \& \ A > 0 \\ -\frac{\pi}{2} & \text{if} \quad C = 0 \ \& \ A < 0 \\ \text{undefined} & \text{if} \quad C = 0 \ \& \ A = 0. \end{cases} \tag{13}$$

The resulting circular variable θ can then be modelled and takes values in $[0, 2\pi)$.

A circular analysis of circumplex data has several benefits: it is more in line with its theoretical definition and it allows us to analyse the blend of the two dimensions Agency and Communion instead of both dimensions separately. This provides us with new insights compared to a separate analysis of the two dimensions that is standard in the literature (see,

[4] The selection of the origin in circumplex data depends on the scaling of the Agency and Communion scores. Agency and Communion are measured on a scale from 1 to 5 and for analysis purposes they are later converted to a scale ranging from -1 to 1. Their respective 0 scores form the origin. Although the scaling influences the average score and spread on the IPC and the intensity, the difference between the individual measurements will be retained (albeit them being different in size). With regards to the intensity, scaling has the same effect as in a standard linear regression. With regards to the location on the IPC (the circular component) rescaling to -1,1 affects the size of the circular (as well as bivariate linear) regression coefficients. This has a positive effect if the original Agency and Communion scores are far from the origin in bivariate space. In that case the locations on the IPC are very concentrated before scaling (small variance). Such data are hard to estimate using circular models as the circular regression coefficients may be very small (high risk of bias, see Cremers, Mainhard, Klugkist (2018a)). Scaling will thus improve estimation. Finally note that scaling is only considered an issue in those instances where cylindrical data is derived from measurements in bivariate space.

e.g., Pennings et al. (2018), Wright et al. (2009) or Wubbels et al. (2006)). There is however one main drawback: when two-dimensional data are converted to the circle we lose some information, namely the length of the two-dimensional vector $(A, C)^t$, *i.e.*, its Euclidean norm $|| (A, C)^t ||$. This length represents the intensity of the interpersonal behavior a teacher shows towards his/her students. In a cylindrical model this intensity (the linear component) can be modeled together with the location of interpersonal behavior of a teacher on the IPC (the circular component). This leads to an improved analysis of interpersonal circumplex data, over either analyzing the two dimensions separately or using a circular model, because we take all information, circular and linear, into account. In the next section we introduce several cylindrical models that can be used to analyze the teacher data. First however we will provide descriptives for the teacher data and conduct a "standard" analysis for these data that we can compare the results from the cylindrical models to.

Data Description

The teacher data was collected between 2010 and 2015 and contains several repeated measures on the IPC of 161 teachers. Measurements were obtained using the QTI and taken in different years and classes. For this paper we only consider one measurement, the first occasion (2010) and largest class if data for multiple classes were available. This results in a sample of 151 teachers. The data includes the location of interpersonal behavior on the IPC (IPC), the circular component, and the intensity of interpersonal behavior (IPC intensity), the linear component. It also includes teachers' self-efficacy (SE) concerning classroom management as covariate that is used to model the IPC and IPC intensity. This means that $\boldsymbol{x}_i = \boldsymbol{z}_i = (1, \text{SE}_i)$ in all cylindrical models except for the Abe-Ley model where $\boldsymbol{z}_i = \text{SE}_i$. In previous research, in psychology and education it has been shown that higher self-efficacy is related to the quality of interpersonal interactions (Locke & Sadler, 2007; Van der Want et al., 2018). After listwise deletion of missings (3 in total, only for the self-efficacy) we have a sample of 148 teachers. Note that we remove the missings here for simplicity, however, in general and especially when the number of missings is larger, the influence of listwise deletion should be investigated and necessary precaution (e.g. imputation of missing values) should be taken. Table 2 shows descriptives for the dataset. Note that we also show the scores on the Agency and Communion component of the IPC before transformation to a circular score. For the circular variable IPC we show sample estimates for the circular mean $\bar{\theta}$ and mean resultant length $\hat{\rho}$. For the linear variables (Agency, Communion, IPC intensity and the covariate SE) we show sample estimates of the linear mean and standard deviation (sd). Figure 6 is a scatterplot showing the relation between the linear and circular component of the teacher data for teachers with low SE (below 1 sd below the mean), average SE (between 1 sd below and 1 sd above the mean) and high SE (above 1 sd above the mean).

[Table 2 about here]

[Figure 6 about here]

Standard Analysis

To perform a standard analysis of the teacher data we fit two linear regression models to the Agency and Communion scores of the teachers separately. In these regression models we incorporate an intercept and the covariate self-efficacy. Coefficients and standard errors from the models are shown in Table 3. From the results we conclude that the effect of self-efficacy is significant for both the Agency and Communion component. An increase of 1 unit on self-efficacy leads to a 0.07 increase in Agency and a 0.09 increase in Communion. Note that in this setup we can only model the effect of self-efficacy on the intensity (size of the score) on the Agency and Communion component. We can not quantify the effect on the location and intensity on the IPC nor can we quantify their dependence.

[Table 3 about here]

Results

In this section we analyze the teacher data with the help of the four cylindrical models from the previous section. We will present the results, posterior estimates and their interpretation for all four models.

Analysis

In the Supplementary Material we have described the starting values for the MCMC procedures for the CL-PN, CL-GPN and GPN-SSN models, hence it remains to specify the starting values for the maximum likelihood based Abe-Ley model: $\eta_0 = 0.9, \eta_1 = 0.9, \nu_0 = 0.9, \nu_1 = 0.9, \kappa = 0.9, \alpha = 0.9, \lambda = 0$. The initial number of iterations for the three MCMC samplers was set to 2000. After convergence checks via traceplots we concluded that some of the parameters of the GPN-SSN model did not converge. Therefore we set the number of iterations of the MCMC models to 20,000 and subtracted a burn-in of 5000 to reach convergence (the Geweke diagnostics show absolute z-scores over 1.96 in 6% of the estimated parameters). Note that we choose the same number of iterations for all three models estimated using MCMC procedures to make their comparison via the PLSL as fair as possible. Lastly, the predictor `SE` was centered before inclusion in the analysis as this allows the intercepts to bear the classical meaning of average behavior.

Tables 4, 5 and 6 show the results for the four cylindrical models that were fit to the teacher data. For the models estimated using MCMC methods, CL-PN, CL-GPN and GPN-SSN, we show descriptives of the posterior of the estimated parameters (posterior mode and lower and upper bound of the 95% highest posterior density (HPD) interval). For the Abe-Ley model we show the maximum likelihood estimates of the parameters. To compare the results of the four models we focus on the following aspects: the estimated average scores (intercept) on the location and intensity on the IPC (1), the effect of self-efficacy on the location and intensity (2), the dependence between the location and intensity (3) and the model fit (4).

[Table 4, 5 and 6 about here]

Average location and intensity on the IPC. The parameters γ_0 in the CL-PN, CL-GPN and Abe-Ley model and the parameter β_{0_y} in the GPN-SSN model inform us about the intensity of interpersonal behavior at the average self-efficacy. For the CL-PN, CL-GPN and GPN-SSN models the parameters are estimated at 0.38, 0.37 and 0.33 respectively and are a direct prediction of the intensity of interpersonal behavior at the average self-efficacy. The estimate for the GPN-SSN model is notably lower and likely to be caused by its skewed distribution for the intensity of interpersonal behavior. In the Abe-Ley model, γ_0 influences the shape parameter of the distribution of the intensity of interpersonal behavior and does not directly estimate the average intensity. Instead we can use the survival function to say something about the probability of having a certain intensity of interpersonal behavior. Figure 7 shows this function for several values of self-efficacy. We look at the survival function at average values of self-efficacy. Note that this function is the average of all survival functions for observations that fall within 1 standard deviation of the mean. The survival function indicates that the probability of having a low intensity of interpersonal behavior is higher than having a high intensity. We however can not make any direct statement about the estimated intensity using the Abe-Ley model.

[Figure 7 about here]

The parameters β_0^I, β_0^{II}, β_0, $\beta_{0_{sI}}$ and $\beta_{0_{sII}}$ inform us about the location on the IPC at the average self-efficacy for the CL-PN, CL-GPN, Abe-Ley and GPN-SSN model respectively. For the CL-PN, CL-GPN and GPN-SSN model we need to combine the estimates for the underlying bivariate components $\{I, II\}$ into one circular estimate using the double arctangent function[5]. Table 7 shows that these circular estimates are similar for the three models at 32.29°, 33.70° and 35.53°. In the Abe-Ley model the location on the IPC at the average self-efficacy is estimated at 0.36 radians or 20.63°.

[5] atan2$(\beta_0^{II}, \beta_0^I)$ or atan2$(\beta_{0_{sII}}, \beta_{0_{sI}})$

[Table 7 about here]

The effect of self-efficacy. The parameters γ_1 in the CL-PN, CL-GPN, Abe-Ley models and β_{1_y} in the GPN-SSN model inform us about the effect of self-efficacy on the intensity of interpersonal behavior. For the CL-PN, CL-GPN and GPN-SSN model the parameters are estimated at 0.03, 0.03 and 0.09 respectively and are a direct estimate of the effect of self-efficacy on the intensity of interpersonal behavior, *i.e.* an increase of 1 unit in self-efficacy leads to an increase of 0.09 units in the intensity of interpersonal behavior according to the GPN-SSN model. These estimates are however quite small and only different from 0 (the HPD interval does not contain 0) in the GPN-SSN model. It is hard to say which of the three models, CL-PN, CL-GPN or GPN-SSN, to use to base our conclusions on. The models CL-GPN and CL-PN fit the linear component best according to the model fit in Table 8. In these models the linear component has a symmetric distribution whereas in the GPN-SSN the distribution of the linear component is skewed. However, the effect of self-efficacy is different from 0 only in the GPN-SSN model which does not seem to match with its lower model fit.

[Table 8 about here]

In the Abe-Ley model, γ_1 influences the shape parameter of the distribution of the intensity of interpersonal behavior and does not directly estimate the effect of self-efficacy. Instead we can use the survival function to say something about the probability of having a certain intensity of interpersonal behavior for different values of self-efficacy. Figure 7 shows this function for low, average and high values of self-efficacy (as defined in Figure 6). This function indicates that the effect of self-efficacy on the intensity of interpersonal behavior is not linear. The probability of having a higher intensity of interpersonal behavior is highest for low self-efficacy and lowest for average self-efficacy.

The parameters β_1^I, β_1^{II}, β_1, $\beta_{1_{sI}}$ and $\beta_{1_{sII}}$ inform us about the effect of self-efficacy on the location on the IPC in the CL-PN, CL-GPN, Abe-Ley and GPN-SSN model respectively.

For the CL-PN and Abe-Ley models we have drawn the circular regression lines for this effect in Figure 8 (see the description of the CL-PN and CL-GPN models for a detailed explanation of circular regression lines). For the CL-PN model the inflection point is indicated with a square in Figure 8. The inflection point for the Abe-Ley model falls outside the bounds of the plot and is therefore not displayed. The slope at the inflection point, b_c, for the CL-PN model is computed by using methods from Cremers et al. (2018b) and is equal to 1.67 (-24.66, 29.33)[6]. The parameter β_1 is the slope at the inflection point for the Abe-Ley model and is equal to -0.03. Even though these slopes are different[7], the regression lines in Figure 8 are quite similar in the data range. Both the regression line of the Abe-Ley model and the CL-PN model show slopes that are not very steep in the range of the data indicating that the effect of self-efficacy on the location on the IPC is not large.

[Figure 8 about here]

In the CL-GPN and GPN-SSN models we cannot compute circular regression coefficients due to the fact that not only the mean vector of the GPN distribution but also the covariance matrix influences the predicted value on the circle. Instead, we will compute posterior predictive distributions for the predicted circular component of individuals scoring the minimum, maximum and median self-efficacy. The modes and 95% HPD intervals of these posterior predictive distributions are $\hat{\theta}_{SE_{min}} = 215.74°(147.36°,\ 44.49°)$, $\hat{\theta}_{SE_{median}} = 25.93°(337.02°,\ 138.59°)$, $\hat{\theta}_{SE_{max}} = 30.86°(8.63°,\ 72.19°)$ for the CL-GPN model. Note that we display the modes and HPD intervals for the posterior predictive distributions on the interval $[0°, 360°)$ and that $44.49° = 404.49°$ due to the periodicity of a circular

[6] Note that this is a linear approximation to the circular regression line representing the slope at a specific point. Therefore it is possible for the HPD interval to be wider than 2π. In this case the interval is much wider and covers 0, indicating there is no evidence for an effect.

[7] Note that the difference in regression lines (and thus inflection point) seems to be influenced by an outlier with a low self-efficacy and IPC value of approximately 160° = -200°.

variable. The posterior mode estimate of 215.74° thus lies within its HPD interval (147.36°, 44.49°). For the GPN-SSN model the modes and 95% HPD intervals of the posterior predictive distributions are $\hat{\theta}_{SE_{min}} = 206.87°(117.12°,\ 72.02°)$, $\hat{\theta}_{SE_{median}} = 24.68°(334.73°,\ 128.27°)$, $\hat{\theta}_{SE_{max}} = 29.81°(0.74°,\ 80.61°)$. For both the CL-GPN and GPN-SSN model the HPD intervals of the mode of the posterior predictive intervals of individuals scoring the minimum, median and maximum self-efficacy overlap. This indicates that the effect of self-efficacy, if there is any, on the location on the IPC shows is not expected to be strong. Had the HPD intervals not overlapped we could have concluded that compared to teachers with a lower self-efficacy increases, the score on the IPC moves counterclockwise.

Dependence between location and intensity on the IPC. The relation between the location and intensity on the IPC in the CL-PN and CL-GPN models is described by the parameters $\gamma_{\cos}$ and $\gamma_{\sin}$. The HPD interval of $\gamma_{\cos}$ does not include 0 for both the CL-PN and CL-GPN models, meaning that the cosine component of the location has an effect on the intensity.

In the teacher data the sine and cosine components have a substantive meaning. This is illustrated in Figure 5. In a unit circle the horizontal axis (Communion) represents the cosine and the vertical axis (Agency) represents the sine of an angle. For the teacher data this means that the Communion (cosine) dimension of the IPC positively effects the intensity of a teachers' interpersonal behavior, in plain words: teachers exhibiting interpersonal behavior types with higher communion scores (e.g., "helpful" and "understanding" in Figure 2) are stronger in their interpersonal behavior.

In the GPN-SSN model the dependence between the location and intensity on the IPC is modelled through the covariances between the linear component and the sine and cosine of the circular component $\sum_{sy_{2,3}}$ and $\sum_{sy_{1,3}}$. Both covariances, $\sum_{sy_{2,3}} = 0.09$ and $\sum_{sy_{1,3}} = 0.23$, are different from zero, but the one of the cosine component, and thus the correlation with the Communion dimension, is larger. This means that teachers scoring both high on Communion and Agency show stronger behavior. Together with the results from the CL-PN and CL-GPN models in the previous paragraph this translates to the conclusion that teachers with higher intensity scores have a location on the IPC that lies between 0° and 90°. To get these scores on the circle both the Agency and the Communion score of a Teacher have to be positive (see (13)). This corresponds to the pattern observed in the teacher data in Figure 5. At an intensity of 0.4 and up we see that the scores on the circle range on average between 0° and 100°.

Model fit. Table 8 shows the values of the PLSL criterion for the linear and circular components of the four cylindrical models fit to the teacher data.

The CL-PN and CL-GPN models have the best out-of-sample predictive performance for the linear component. They show roughly the same performance because they model the linear component in the same way. We should note that even though the predictive performance of the Abe-Ley model for the linear component is worst on average, the standard deviation of the cross-validation estimates is rather large. This means that in some samples, the Abe-Ley model shows a lower PLSL value than the average of 25.49.

The Abe-Ley model has the best out-of-sample predictive performance for the circular component. This would suggest that for the circular variable a slightly skewed distribution fits best. However, both the GPN-SSN and the CL-GPN models fit much worse even though the distribution for the circular component in these models can also take a skewed shape. It should be noted that the standard deviation of the cross-validation estimates is rather large for the Abe-Ley and the CL-GPN model. It is possible that these large standard deviations for the PLSL are caused by the fact that they are computed for a relatively small sample size, but this does not explain why the PLSL has a large standard deviation for only a few cylindrical models and not for all.

In this situation, where one model fits the linear component best and another one fits the circular component best, it is hard to determine which model we should choose. In this case the results for the CL-PN /CL-GPN and Abe-Ley model are quite different regarding the effect of self-efficacy on the linear component (intensity of interpersonal behavior).

Because the Abe-Ley fit for the linear part is worst we would choose to trust the results for the CL-PN and CL-GPN models here. For the circular part however the results of the CL-PN/CL-GPN models do not differ as much from the Abe-Ley model and we reach the same conclusion for both models, namely that the effect of self-efficacy on location on the IPC is not very strong. Therefore we would prefer the CL-PN/CL-GPN models in this case because where it matters in terms of interpretation (the linear part) they show better fit.

Discussion

In this paper we modified four models for cylindrical data in such a way that they include a regression of both the linear and circular component onto a set of covariates. Subsequently we have shown how these four methods can be used to analyze a dataset on the interpersonal behavior of teachers. In this final section we will first comment on what researchers can gain by using cylindrical models for the teacher data. Subsequently we will comment on the differences between the cylindrical models that were introduced in this paper.

Concerning the teacher data, the advantage of cylindrical data analysis is that we were able to analyze the information about the location and intensity on the IPC simultaneously. In previous research, the two components of the interpersonal circumplex (*i.e.*, Agency and Communion) were analyzed separately. Such an approach only provides information about the intensity of a teachers' score on Agency and Communion and the information about the combination of Agency and Communion, which describes the location on the IPC, gets lost as was observed in the Standard Analysis section. A first solution to include both dimensions as a circular variable in data analysis was described by Cremers et al. (2018a). A downside of that analysis was that information about the intensity could not be retained. In the present study, we have shown how using cylindrical models can simultaneously model the information about the location and intensity on the IPC and how these are influenced by teachers' self-efficacy in classroom management. Although we do not find any strong effects of self-efficacy on either the location and intensity[8], the four cylindrical models do provide a way of analyzing and interpreting this effect. This is beneficial for future research in which we may want to investigate the effect of further covariates on data from the circumplex.

Furthermore, in addition to being able to assess the influence of covariates, the cylindrical models also provide information about the dependence between the location and intensity on the IPC. This is a second advantage over the standard analysis in which it is impossible to quantify the relation between intensity and location. We found that stronger behavior is associated with higher scores on the Communion and in some models also the Agency dimension. This implies that teachers whose interpersonal behavior ranges between $0°$ and $90°$ on the IPC, the "helpful" and "directing" subtypes, are stronger in their behavior than teachers of the other subtypes.

As mentioned in the introduction, data from the interpersonal circumplex is not the only type of cylindrical data that occurs in psychology. The methods presented in this paper are also of use for research on human navigation and eye-tracking research. Furthermore, even though cylindrical models are already used in fields outside of psychology, the addition of a regression structure to the models is of use in these fields as well.

Out of the four cylindrical models investigated in this paper, the results from CL-PN and Abe-Ley models are most straightforward to interpret. In the CL-GPN and GPN-SSN models the interpretation of the parameters of the circular outcome component is more complex, if at all possible. This is caused by the fact that in addition to the mean vector the covariance matrix of the GPN distribution affects the location of the circular data, making it difficult to compute regression coefficients on the circle. Wang and Gelfand (2013) state that Monte Carlo integration can be used to compute a circular mean and variance for the GPN distribution. In future research, this solution might be applied to the methods of Cremers et al. (2018b) in order to compute circular coefficients for GPN models.

The GPN-SSN model is more flexible compared to the other three cylindrical models

[8] similar to the standard analysis

that were investigated in this paper. Multiple linear and circular components can be included and we can thus apply the model to multivariate cylindrical data. In addition the GPN-SSN, the CL-GPN and CL-PN models are extendable to a mixed-effects structure and can thus also be fit to longitudinal data (see Nuñez-Antonio and Gutiérrez-Peña (2014) and Hernandez-Stumpfhauser et al. (2016) for hierarchical/mixed-effects models for the PN and GPN distributions respectively). For the Abe-Ley model this may also be possible but has not been done in previous research for the conditional distribution of its circular component (sine-skewed von Mises). Concerning asymmetry, both the GPN-SSN as well as the Abe-Ley model allow for non-symmetrical shapes of the distributions of both the linear and circular component, while the CL-GPN model permits an asymmetric circular component.

The four cylindrical models that were modified to the regression context in this paper are not the only cylindrical distributions available from the literature. Other interesting cylindrical distributions have been introduced by Fernández-Durán (2007), Kato and Shimizu (2008) and Sugasawa (2015) (for more references we refer to Chapter 2 of Ley and Verdebout (2017)). In the present study we have decided not to include these distributions for reasons of space, complexity of the models and ease of implementing a regression structure. In future research however it would be interesting to investigate other types of cylindrical distributions as well in order to compare the interpretability, flexibility and model fit to the models developed in the present study.

References

Abe, T., & Ley, C. (2017). A tractable, parsimonious and flexible model for cylindrical data, with applications. *Econometrics and Statistics*, *4*, 91–104. doi:10.1016/j.ecosta.2016.04.001

Chrastil, E. R., & Warren, W. H. (2017). Rotational error in path integration: Encoding and execution errors in angle reproduction. *Experimental Brain Research*, *235*(6), 1885–1897. doi:10.1007/s00221-017-4910-y

Claessens, L. C. (2016). *Be on my side i'll be on your side : Teachers' perceptions of teacher–student relationships* (PhD thesis).

Cremers, J., Mainhard, M. T., & Klugkist, I. (2018a). Assessing a Bayesian embedding approach to circular regression models. *Methodology*, *14*(2), 69–81.

Cremers, J., Mulder, K. T., & Klugkist, I. (2018b). Circular interpretation of regression coefficients. *British Journal of Mathematical and Statistical Psychology*, *71*(1), 75–95. doi:10.1111/bmsp.12108

Fernández-Durán, J. (2007). Models for circular–linear and circular–circular data constructed from circular distributions based on nonnegative trigonometric sums. *Biometrics*, *63*(2), 579–585. doi:10.1111/j.1541-0420.2006.00716.x

Fisher, N. I. (1995). *Statistical analysis of circular data.* Cambridge: Cambridge University Press.

García-Portugués, E., Barros, A. M., Crujeiras, R. M., González-Manteiga, W., & Pereira, J. (2014). A test for directional-linear independence, with applications to wildfire orientation and size. *Stochastic Environmental Research and Risk Assessment*, *28*(5), 1261–1275. doi:10.1007/s00477-013-0819-6

García-Portugués, E., Crujeiras, R. M., & González-Manteiga, W. (2013). Exploring wind direction and SO2 concentration by circular–linear density estimation. *Stochastic Environmental Research and Risk Assessment*, *27*(5), 1055–1067. doi:10.1007/s00477-012-0642-5

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378. doi:10.1198/016214506000001437

Gurtman, M. B. (2009). Exploring personality with the interpersonal circumplex. *Social and Personality Psychology Compass*, *3*(4), 601–619. doi:10.1111/j.1751-9004.2009.00172.x

Gurtman, M. B. (2011). Handbook of interpersonal psychology. In L. M. Horowitz & S. Strack (Eds.), (pp. 299–311). New York: Wiley.

Hernandez-Stumpfhauser, D., Breidt, F. J., & Van der Woerd, M. J. (2016). The general projected normal distribution of arbitrary dimension: Modeling and Bayesian inference. *Bayesian Analysis*, *12*(1), 113–133. doi:10.1214/15-BA989

Horowitz, L. M., & Strack, S. (2011). *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions*. Hoboken, NJ: John Wiley & Sons.

Jammalamadaka, S. R., & Sengupta, A. (2001). *Topics in circular statistics* (Vol. 5). World Scientific.

Johnson, R. A., & Wehrly, T. E. (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, *73*(363), 602–606.

Kato, S., & Shimizu, K. (2008). Dependent models for observations which include angular ones. *Journal of Statistical Planning and Inference*, *138*(11), 3538–3549. doi:10.1016/j.jspi.2006.12.009

Lagona, F., Picone, M., Maruotti, A., & Cosoli, S. (2015). A hidden Markov approach to the analysis of space–time environmental data with linear and circular components. *Stochastic Environmental Research and Risk Assessment*, *29*(2), 397–409. doi:10.1007/s00477-014-0919-y

Leary, T. (1957). *An interpersonal diagnosis of personality.* New York: Ronald Press Company.

Ley, C., & Verdebout, T. (2017). *Modern directional statistics.* Chapman & Hall/CRC Press. Boca Raton, FL.

Locke, K. D., & Sadler, P. (2007). Self-efficacy, values, and complementarity in dyadic interactions: Integrating interpersonal and social-cognitive theory. *Personality and Social Psychology Bulletin*, *33*, 94–109. doi:10.1177/0146167206293375

Mardia, K. V., & Jupp, P. E. (2000). *Directional statistics* (Vol. 494). Chichester, England: Wiley.

Mardia, K. V., & Sutton, T. W. (1978). A model for cylindrical variables with applications. *Journal of the Royal Statistical Society Series B (Methodological)*, *40*(2), 229–233.

Mastrantonio, G. (2018). The joint projected normal and skew-normal: A distribution for poly-cylindrical data. *Journal of Multivariate Analysis*, *165*, 14–26. doi:10.1016/j.jmva.2017.11.006

Mastrantonio, G., Maruotti, A., & Jona-Lasinio, G. (2015). Bayesian hidden Markov modelling using circular-linear general projected normal distribution. *Environmetrics*, *26*(2), 145–158. doi:10.1002/env.2326

Nuñez-Antonio, G., & Gutiérrez-Peña, E. (2014). A Bayesian model for longitudinal circular data based on the projected normal distribution. *Computational Statistics & Data Analysis*, *71*, 506–519. doi:10.1016/j.csda.2012.07.025

Nuñez-Antonio, G., Gutiérrez-Peña, E., & Escarela, G. (2011). A Bayesian regression model for circular data based on the projected normal distribution. *Statistical Modelling*, *11*(3), 185–201. doi:10.1177/1471082X1001100301

Pennings, H. J. M. (2017a). *Interpersonal dynamics in teacher-student interactions and relationships* (PhD thesis). Enschede: Ipskamp drukkers.

Pennings, H. J. M. (2017b). Using a complexity approach to study the interpersonal dynamics in teacher-student interactions: A case study of two teachers. *Complicity: An International Journal of Complexity and Education*, *14*(2), 88–103.

Pennings, H. J. M., Brekelmans, M., Sadler, P., Claessens, L. C., Van der Want, A. C., & Van Tartwijk, J. (2018). Interpersonal adaptation in teacher-student interaction. *Learning and Instruction*, *55*, 41–57. doi:10.1016/j.learninstruc.2017.09.005

Plutchik, R. (1997). Circumplex models of personality and emotions. In C. Plutchik Robert. (Ed.), (pp. 17–45). Washington, DC, US: American Psychological Association.

Rayner, K. (2009). The 35th Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. doi:10.1080/17470210902816461

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. doi:10.1037/h0077714

Sahu, S. K., Dey, D. K., & Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *Canadian Journal of Statistics*, *31*(2), 129–150. doi:10.2307/3316064

Schwartz, S. H. (1992). Advances in experimental social psychology. In M. Zanna (Ed.), (Vol. 25, pp. 1–65). San Diego, CA: Academic Press.

Sugasawa, S., S. (2015). *A flexible family of distributions on the cylinder.* Retrieved from arXiv: 1501.06332v2

Wang, F., & Gelfand, A. E. (2013). Directional data analysis under the general projected normal distribution. *Statistical Methodology*, *10*(1), 113–127. doi:10.1016/j.stamet.2012.07.005

Van der Want, A. C. (2015). *Teachers' interpersonal role identity.* (PhD thesis).

Van der Want, A. C., Den Brok, P., Beijaard, D., Brekelmans, M., Claessens, L. C., & Pennings, H. J. M. (2018). The relation between teachers' interpersonal role identity and their self-efficacy, burnout and work engagement. *Professional Development in Education*, Advance Online Publication. doi:10.1080/19415257.2018.1511453

Wiggins, J. (1996). An informal history of the interpersonal circumplex tradition. *Journal of Personality Assessment*, *66*(2), 217–233.

Wright, A. G., Pincus, A. L., Conroy, D. E., & Hilsenroth, M. J. (2009). Integrating methods to optimize circumplex description and comparison of groups. *Journal of Personality Assessment*, *91*(4), 311–322. doi:10.1080/00223890902935696

Wubbels, T., Brekelmans, M., Den Brok, P., & Van Tartwijk, J. (2006). Handbook of classroom management: Research practice and contemporary issues. In C. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary* *issues* (pp. 1161–1191). Malwah, NJ: Lawrence Erlbaum Associates.

Table 1

Comparison of the four cylindrical regression models

Aspect	CL-PN	CL-GPN	Abe-Ley	GPN-SSN
Θ				
Distribution	PN	GPN	Sine-skewed vM	GPN
Domain	$[0, 2\pi)$	$[0, 2\pi)$	$[0, 2\pi)$	$[0, 2\pi)$
Shape	symmetric, unimodal	asymmetric, multimodal	asymmetric, unimodal	asymmetric, multimodal
Y				
Distribution	Normal	Normal	Weibull	skewed-Normal
Domain	$(-\infty, +\infty)$	$(-\infty, +\infty)$	$(0, +\infty)$	$(-\infty, +\infty)$
Shape	symmetric, unimodal	symmetric, unimodal	asymmetric, unimodal	asymmetric, unimodal
Θ-Y dependence				
	Y regressed on $\sin(\Theta)$ and $\cos(\Theta)$	Y regressed on $\sin(\Theta)$ and $\cos(\Theta)$	circular concentration κ and linear scale α	covariance matrix

Note: PN and GPN refer to the projected normal and general projected normal distribution.

vM refers to the von-Mises distribution

Table 2

Descriptives for the teacher dataset.

Variable	mean/$\bar{\theta}$	sd/$\hat{\rho}$	Range	Type
Agency	0.19	0.16	-0.33 - 0.49	Linear
Communion	0.30	0.24	-0.58 - 0.77	Linear
IPC	33.22°	0.76	-	Circular
intensity IPC	0.43	0.15	0.08 - 0.80	Linear
SE	5.04	1.00	1.5 - 7.0	Linear

Note: For the circular variable IPC we show sample estimates for the circular mean $\bar{\theta}$ and mean resultant length $\hat{\rho}$. For the linear variable we show the sample mean, standard deviation and range.

Table 3

Regression coefficients and standard errors for the standard analysis of the teacher dataset.

	Agency	Communion
Intercept	0.19 (0.01)	0.30 (0.02)
SE	0.07 (0.01)	0.09 (0.02)

Note: The intercepts indicate the average Agency and Communion. The coefficient for SE indicates the effect of self-efficacy on Agency and Communion.

Table 4

Results, cross-validation mean and standard deviation, for the modified CL-PN and CL-GPN models

Parameter	CL-PN			CL-GPN		
	Mode	HPD LB	HPD UB	Mode	HPD LB	HPD UB
β_0^I	1.76 (0.09)	1.50 (0.07)	2.03 (0.09)	2.43 (0.12)	1.91 (0.10)	3.05 (0.17)
β_1^I	0.65 (0.07)	0.42 (0.06)	0.90 (0.08)	0.84 (0.11)	0.45 (0.09)	1.29 (0.15)
β_0^{II}	1.15 (0.05)	0.92 (0.04)	1.37 (0.04)	1.47 (0.05)	1.16 (0.04)	1.78 (0.05)
β_1^{II}	0.58 (0.03)	0.38 (0.04)	0.79 (0.04)	0.70 (0.06)	0.47 (0.05)	0.96 (0.08)
γ_0	0.38 (0.01)	0.31 (0.01)	0.44 (0.01)	0.37 (0.01)	0.31 (0.01)	0.42 (0.01)
γ_{cos}	0.04 (0.00)	0.01 (0.00)	0.06 (0.00)	0.03 (0.00)	0.01 (0.00)	0.04 (0.00)
γ_{sin}	-0.01 (0.00)	-0.04 (0.00)	0.02 (0.00)	-0.00 (0.00)	-0.03 (0.00)	0.03 (0.00)
γ_1	0.03 (0.01)	-0.00 (0.00)	0.07 (0.01)	0.03 (0.00)	-0.00 (0.00)	0.06 (0.00)
σ	0.14 (0.00)	0.12 (0.00)	0.16 (0.00)	0.14 (0.00)	0.12 (0.00)	0.16 (0.00)
$\sum_{1,1}$	NA (NA)	NA (NA)	NA (NA)	3.04 (0.29)	1.85 (0.13)	5.00 (0.41)
$\sum_{1,2}$	NA (NA)	NA (NA)	NA (NA)	0.47 (0.12)	0.12 (0.12)	0.80 (0.10)
$\sum_{2,2}$	NA (NA)	NA (NA)	NA (NA)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)

Note: β_0^I, β_0^{II} and γ_0 inform us about the location and intensity on the IPC at the average self-efficacy. β_1^I, β_1^{II} and γ_1 inform us about the effect of self-efficacy on the location and intensity on the IPC. γ_{cos} and γ_{sin} inform us about the dependence between the location and intensity on the IPC. $\sum_{1,1}$, $\sum_{1,2}$ and $\sum_{2,2}$ are elements of the variance-covariance matrix of the location on the IPC in the CL-GPN model and σ is the error standard deviation of the intensity of interpersonal behavior.

Table 5

Results, cross-validation mean and standard deviation (SD), for the modified Abe-Ley model

	β_0	β_1	γ_0	γ_1	α	κ	λ
Mean	0.36	-0.03	1.17	0.04	3.66	1.51	0.70
SD	0.02	0.01	0.02	0.02	0.12	0.08	0.05

Note: β_0 and γ_0 inform us about the location and intensity on the IPC at the average self-efficacy. β_1 and γ_1 inform us about the effect of self-efficacy on the location and intensity on the IPC. α is the shape parameter of the distribution of the location on the IPC. κ and λ respectively are the concentration and skewness parameters for the distribution of the location on the IPC.

Table 6

Results, cross-validation mean and standard deviation, for the GPN-SSN model

Parameter	Unconstrained			Constrained		
	Mode	HPD LB	HPD UB	Mode	HPD LB	HPD UB
$\beta_{0_s^I}$	0.30 (0.01)	0.26 (0.01)	0.34 (0.01)	2.11 (0.11)	1.75 (0.09)	2.50 (0.11)
$\beta_{0_s^{II}}$	0.19 (0.00)	0.17 (0.01)	0.21 (0.00)	1.34 (0.06)	1.10 (0.05)	1.57 (0.06)
β_{0_y}	0.33 (0.01)	0.30 (0.30)	0.36 (0.01)	0.33 (0.01)	0.30 (0.01)	0.36 (0.01)
$\beta_{1_s^I}$	0.09 (0.01)	0.05 (0.01)	0.13 (0.01)	0.60 (0.06)	0.33 (0.05)	0.90 (0.06)
$\beta_{1_s^{II}}$	0.07 (0.00)	0.04 (0.00)	0.09 (0.01)	0.48 (0.03)	0.30 (0.04)	0.66 (0.04)
β_{1_y}	0.09 (0.01)	0.06 (0.06)	0.12 (0.01)	0.09 (0.01)	0.06 (0.01)	0.12 (0.01)
$\sum_{s_{1,1}}$	0.05 (0.00)	0.04 (0.00)	0.06 (0.00)	2.44 (0.15)	1.72 (0.07)	3.46 (0.14)
$\sum_{s_{2,2}}$	0.02 (0.00)	0.02 (0.00)	0.03 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
$\sum_{y_{3,3}}$	0.03 (0.00)	0.02 (0.02)	0.04 (0.00)	0.03 (0.00)	0.02 (0.00)	0.04 (0.00)
$\sum_{s_{1,2}}$	0.00 (0.00)	-0.00 (0.00)	0.01 (0.00)	0.08 (0.06)	-0.20 (0.06)	0.34 (0.06)
$\sum_{sy_{1,3}}$	0.03 (0.00)	0.02 (0.00)	0.04 (0.00)	0.23 (0.01)	0.17 (0.00)	0.32 (0.01)
$\sum_{sy_{2,3}}$	0.01 (0.00)	0.01 (0.01)	0.02 (0.00)	0.09 (0.01)	0.06 (0.01)	0.12 (0.01)
λ	0.16 (0.01)	0.14 (0.01)	0.18 (0.01)	0.16 (0.01)	0.14 (0.01)	0.18 (0.01)

Note: $\beta_{0_s^I}$, $\beta_{0_s^{II}}$ and β_{0_y} inform us about the location and intensity on the IPC at the average self-efficacy. $\beta_{1_s^I}$, $\beta_{1_s^{II}}$ and β_{1_y} inform us about the effect of self-efficacy on the location and intensity on the IPC. $\sum_{s_{1,1}}$, $\sum_{s_{1,2}}$, $\sum_{s_{2,2}}$, $\sum_{y_{3,3}}$, $\sum_{sy_{1,3}}$, and $\sum_{sy_{2,3}}$ are elements of the variance-covariance matrix of which $\sum_{sy_{1,3}}$ and $\sum_{sy_{2,3}}$ inform us about the dependence between the location and intensity on the IPC.

λ is the skewness parameter of the distribution of the intensity of interpersonal behavior.

Table 7

Posterior estimates (in degrees) for the circular mean (at SE = 0) in the CL-PN, CL-GPN and GPN-SSN models

	Mode	HPD LB	HPD UB
CL-PN	32.29	24.81	39.71
CL-GPN	33.70	26.72	41.15
GPN-SSN	35.53	28.40	43.30

Note that these means are based on their posterior predictive distribution following (Wang and Gelfand, 2013)

Table 8

PLSL criteria, cross-validation mean and standard deviation, for the circular and linear component in the four cylindrical models

Model	Circular		Linear	
	mean	sd	mean	sd
CL-PN	82.96	(9.47)	-17.65	(3.70)
CL-GPN	78.21	(14.53)	-18.30	(3.00)
Abe-Ley	31.97	(22.07)	25.49	(17.46)
GPN-SSN	107.10	(10.52)	-2.37	(7.01)

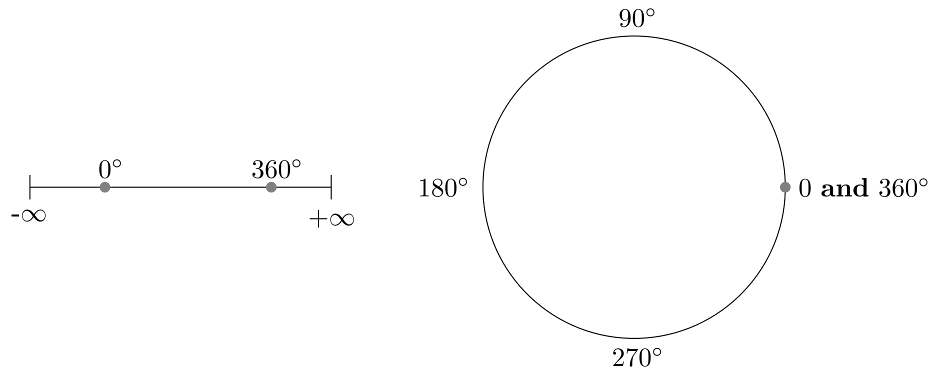

Figure 1. The difference between a linear scale (left) and a circular scale (right).

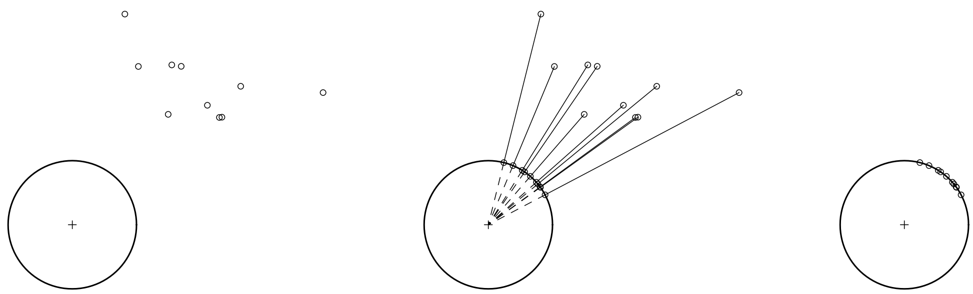

Figure 2. Plot showing the projection of datapoints in bivariate space, $\boldsymbol{S}$, (left) to the circle (right). The lines connecting the bivariate datapoints to the circular datapoints represent the euclidean norms of the bivariate datapoints, realizations of the random variable R.

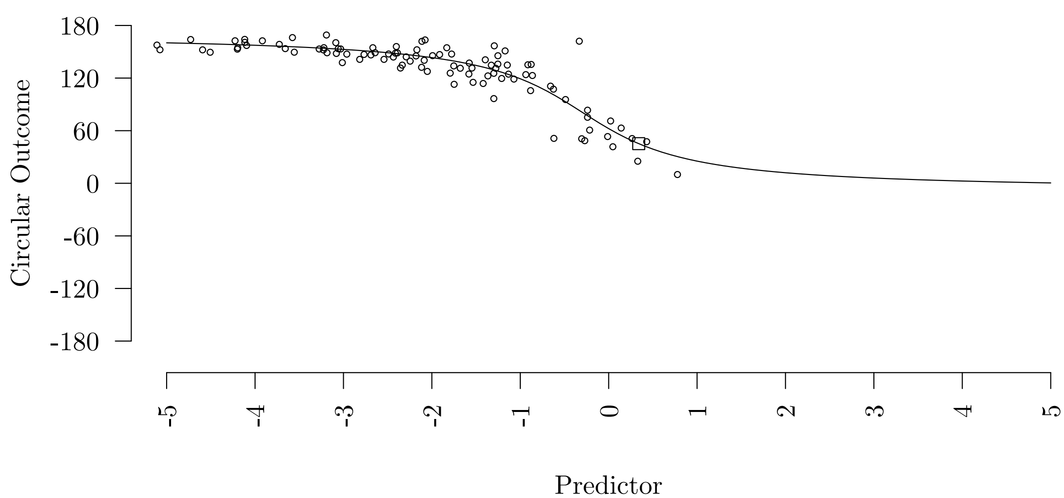

Figure 3. Circular regression line for the relation between a covariate and a circular component with the simulated data the regression line was fit to. The square indicates the inflection point of the regression line.

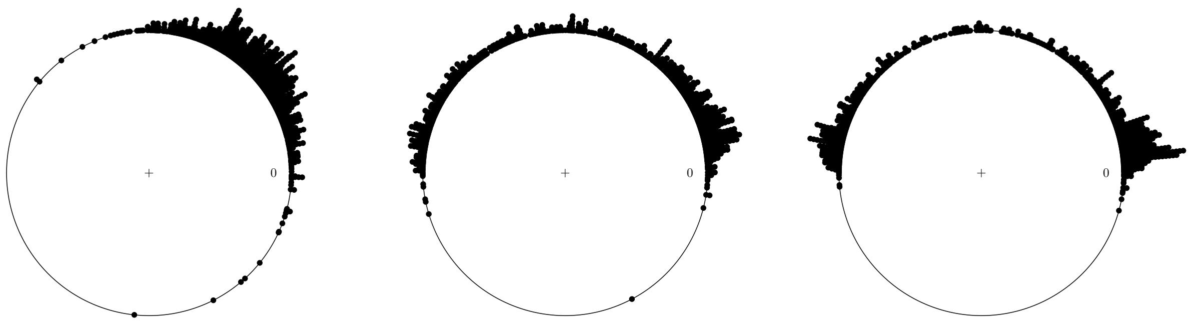

Figure 4. Plot showing 1000 samples from three GPN distributions with mean vector $(2, 2)$ and $\tau = 1, 5, 10$ and $\xi = 0, 0.9, -0.5$ from left to right

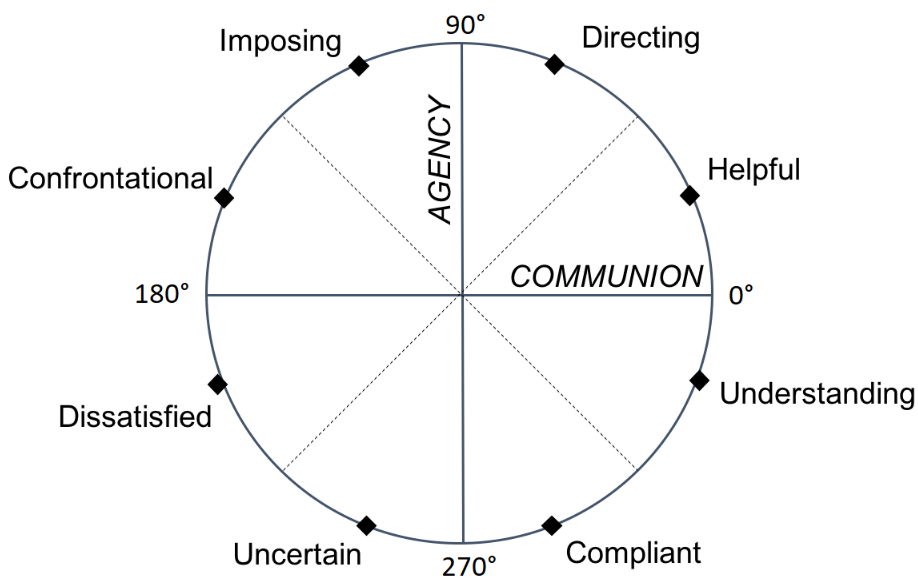

Figure 5. The interpersonal circle for teachers (IPC-T). The words presented in the circumference of the circle are anchor words to describe the type of behavior located in each part of the IPC.

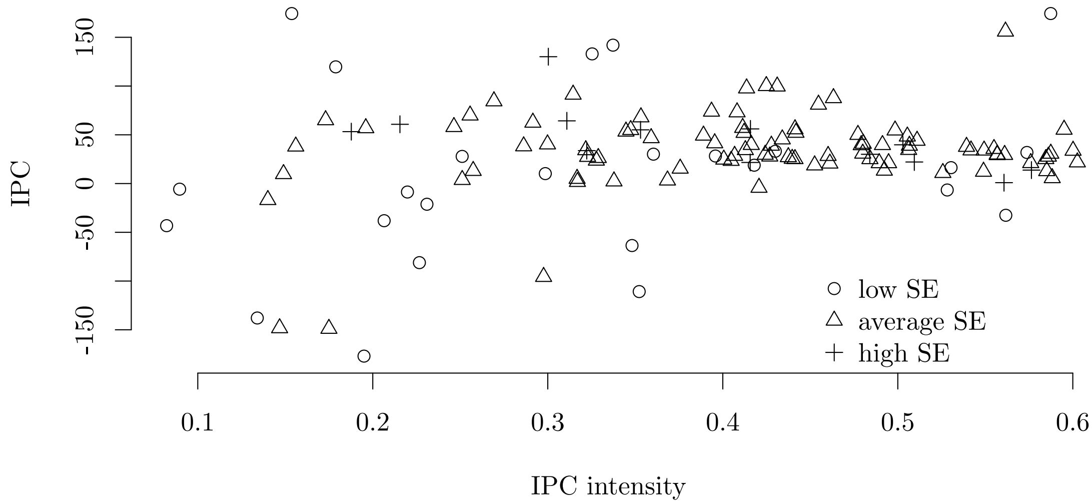

Figure 6. Plot showing the relation between the linear and circular outcome component (in degrees) of the teacher data.

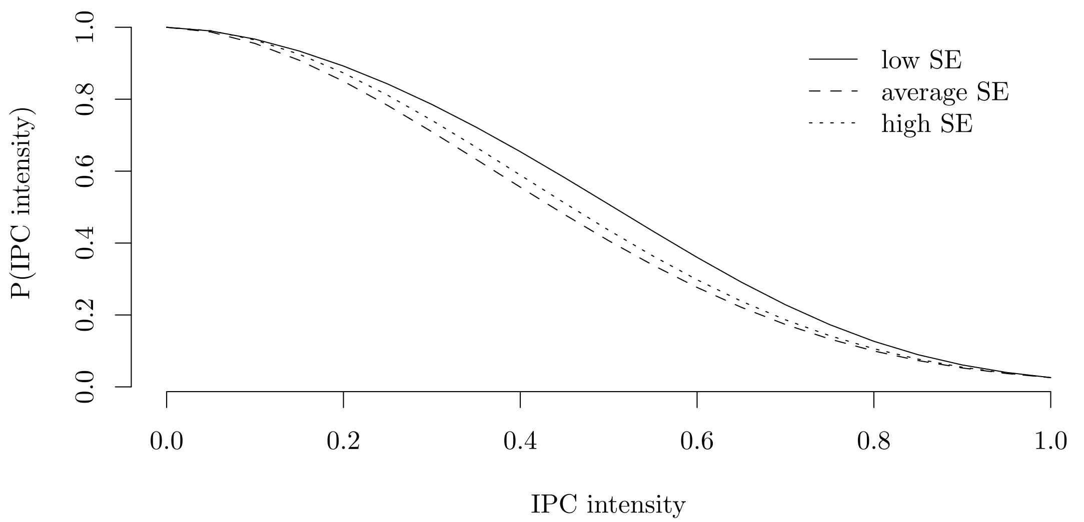

Figure 7. Plot showing the probability of having a particular intensity of interpersonal behavior (survival plot) for the minimum, mean and maximum self-efficacy in the data.

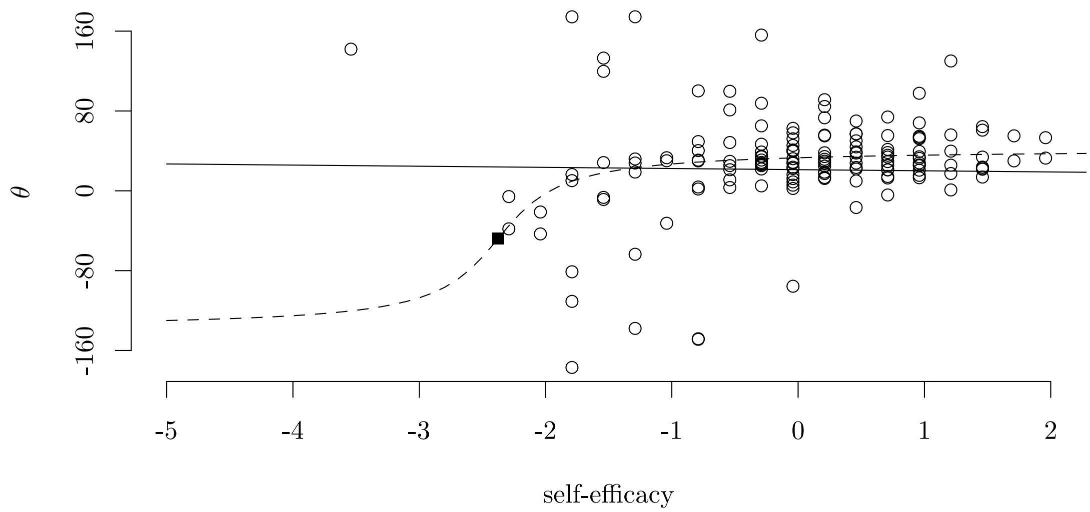

Figure 8. Plot showing circular regression lines for the effect of self-efficacy as predicted by the Abe-Ley model (solid line) and CL-PN model (dashed line). The black square indicates the inflection point of the circular regression line for the CL-PN model.