Chill Portions and yield using actual data

Katja Schiffers[a], *Cory Whitney*[a], *Eduardo Fernandez*[a] *and Eike Luedeling*[a]
[a]*INRES-Horticultural Sciences, University of Bonn, Auf dem Huegel 6, 53121 Bonn, Germany*

Data is often limited for assessing the relationship between temperatures and yield. Despite this lack of data we may still be able to make assessments and produce useful projections for farmers and decision makers. With the right tools this data limitation does not need to hinder our abilities to assess the relationships between temperature and yield. For coarse assessments a lot of data may not be necessary. Here we show a method whereby we can embrace the inherint uncertainty in such assessments to overcome the need for preciseness. We use the pasitR package (Schiffers et al., 2018) in the R programming language (R Core Team, 2019). We show a potential method for dealing with important but also necesarily uncertain relationships in model forecasts.

Yield and chill data

In this file we applied the procedure for estimating yield as a function of chill accumulation. We used data provided by the experimental orchard of the School of Agronomy at the Pontificia Universidad Catolica de Valparaiso. The data set contains yield records (in tons per hectare) for 8 seasons (2010 to 2017) for two sweet cherry cultivars (Lapins and Brooks).

Year	Variety	Yield
2010	Lapins	16.387600
2011	Lapins	11.401600
2012	Lapins	1.599200
2013	Lapins	13.521200
2014	Lapins	21.648480
2015	Lapins	9.413200
2016	Lapins	24.974440
2017	Lapins	8.515682
2010	Brooks	6.412400
2011	Brooks	1.296000
2012	Brooks	1.032000
2013	Brooks	3.396800
2014	Brooks	6.872400
2015	Brooks	2.887160
2016	Brooks	9.217320
2017	Brooks	9.892581

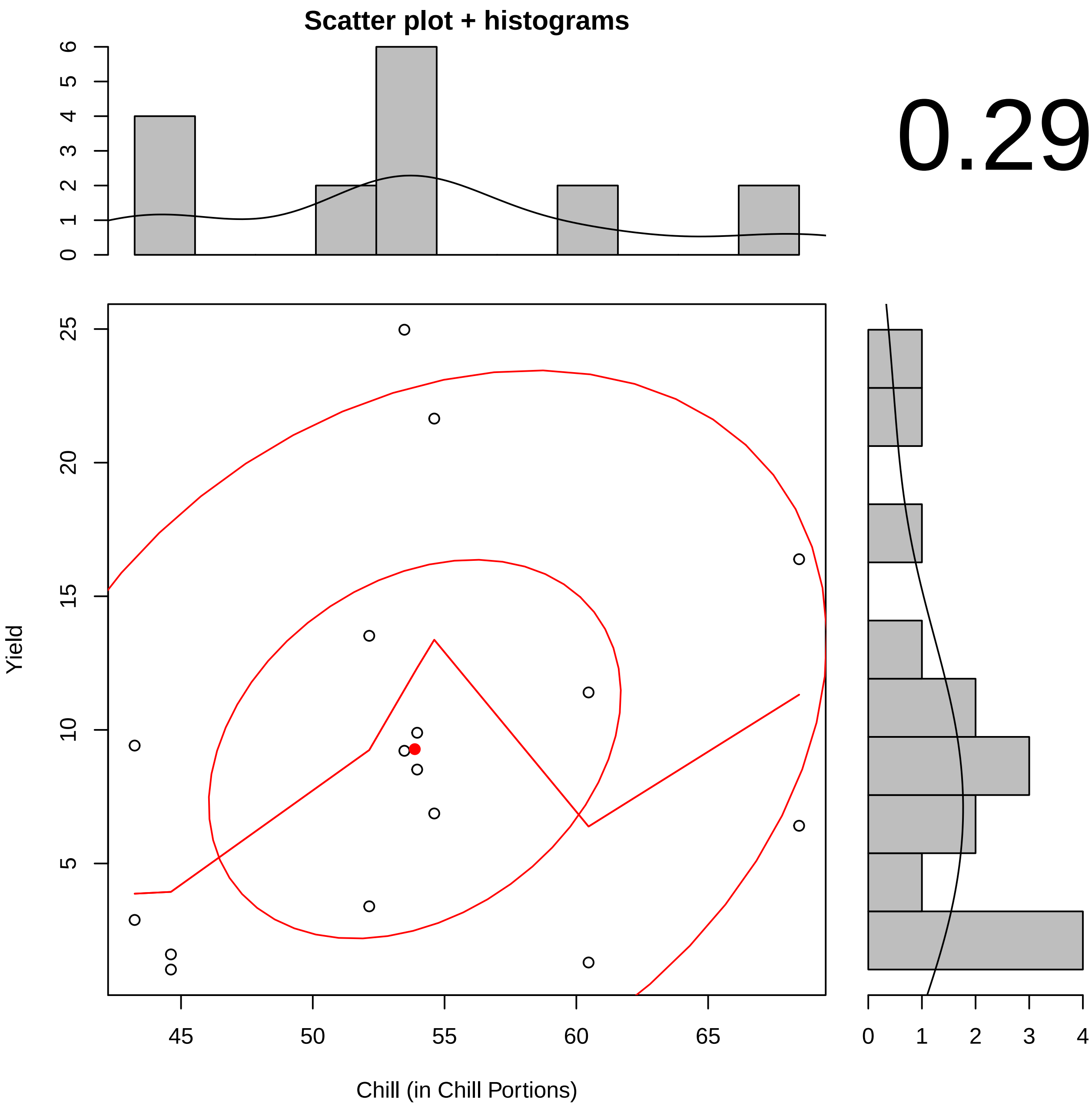
Weather data was obtained from a previous project (chill models comparison) from a weather station placed in the orchard.

YEARMODA	Weather_Station	Year	Month	Day	JDay	Tmin	Tmax
20100101	Quillota	2010	1	1	1	6.2	31.0
20100102	Quillota	2010	1	2	2	7.6	29.4
20100103	Quillota	2010	1	3	3	12.2	23.2
20100104	Quillota	2010	1	4	4	8.0	24.0
20100105	Quillota	2010	1	5	5	8.0	26.5
20100106	Quillota	2010	1	6	6	9.0	27.0
20100107	Quillota	2010	1	7	7	8.5	24.0
20100108	Quillota	2010	1	8	8	7.0	28.0
20100109	Quillota	2010	1	9	9	8.0	28.0
20100110	Quillota	2010	1	10	10	8.0	24.0

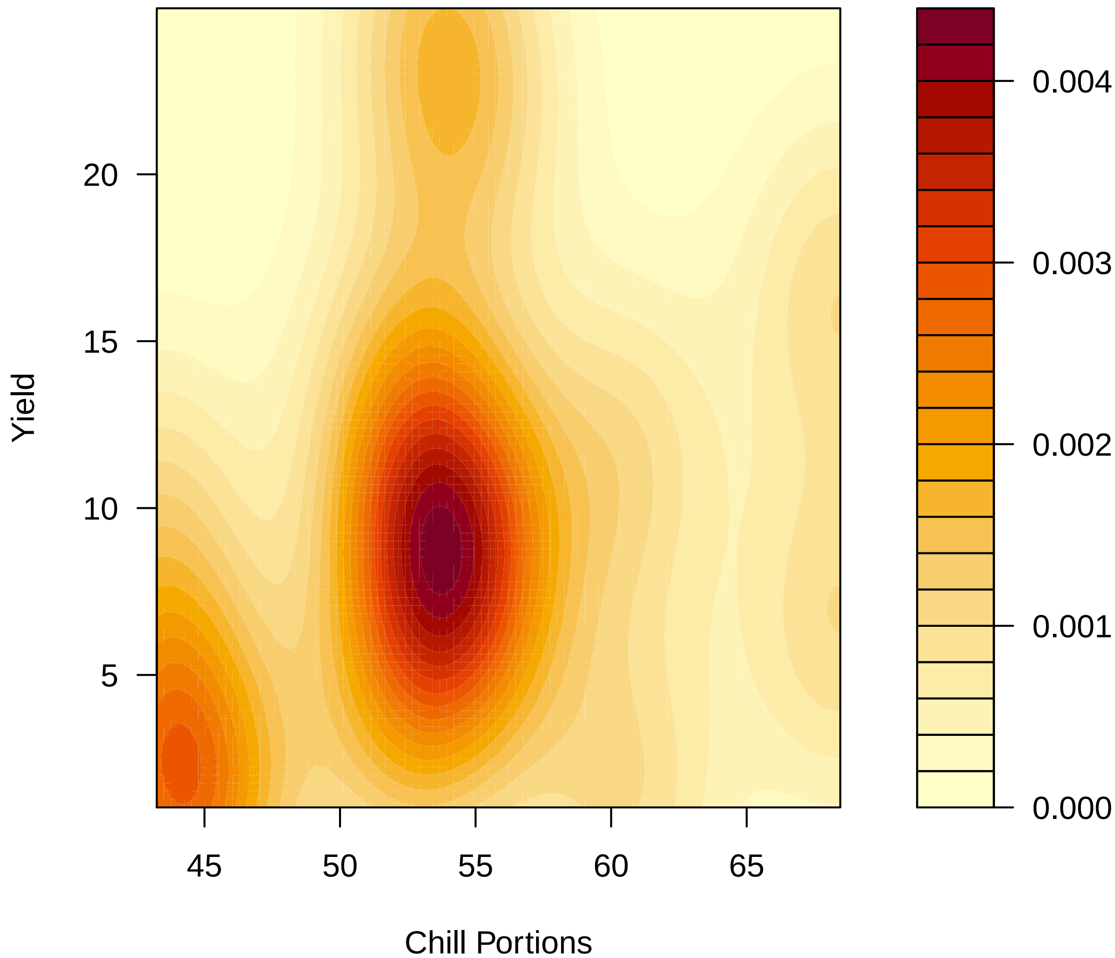
Using `tempResponse_daily_list()` in the `chillR()` package (Luedeling, 2019) compute the chill accumulation for each season. In this example we defined the chilling season as the period between 1st of May and 31st of August.

Season	End_year	Season_days	Data_days	Perc_complete	Chill_Portions
2009/2010	2010	123	123	100	68.45256
2010/2011	2011	123	123	100	60.46397
2011/2012	2012	123	123	100	44.61716
2012/2013	2013	123	123	100	52.14200
2013/2014	2014	123	123	100	54.60832
2014/2015	2015	123	123	100	43.24068
2015/2016	2016	123	123	100	53.47477
2016/2017	2017	123	123	100	53.95999

Using `chillscatter()` create a Chill Portions (x) and yield (y) scatter plot with associated and estimated densities with loess smooth linear fits density curves. Plot made with the `scatter.hist()` function in the `plyr()` package (Wickham, 2019). The following scatter plots show the outputs for Lapins and Brooks varieties.



Use `chillkernel()` or another type of graphical representation of the same relationship. `chillkernel()` performs a two-dimensional kernel density estimation for yield and chill using the `kde2d()` function in the `MASS()` package (Ripley, 2019). The function produces a matrix of the estimated density (z) of yield (y) and chill (x). As the density function restricts the shape of the kernel to a bivariate normal kernel, it looks slightly different compared to the scatter plot estimates above.

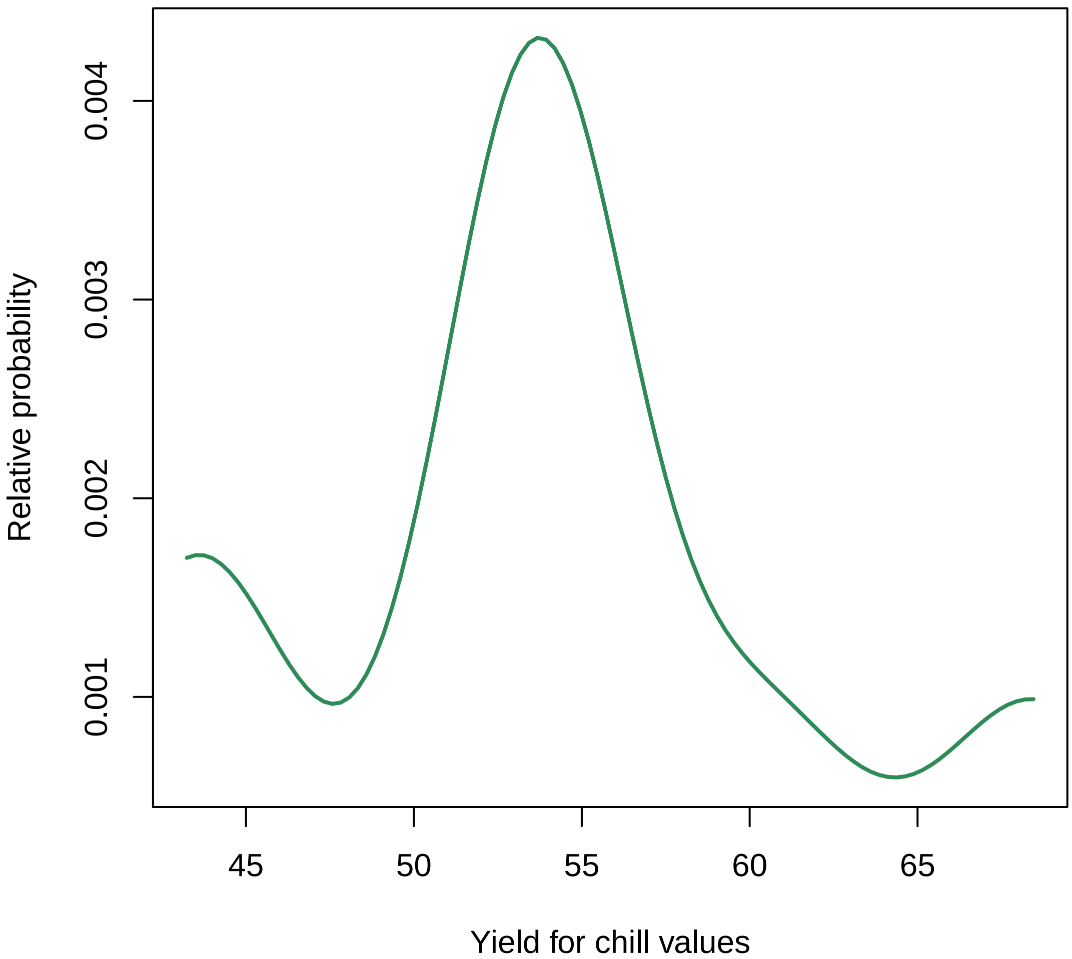


`chillkernel()` shows a density surface plot of Chill Portions (x) and yield (y). The legend shows the value for the estimated density (z). The plot is made with the `filled.contour()` function of the `graphics()` package (R Core Team, 2019).

In `chillkernel()` the density (z) over the entire plot integrates to one, and therefore represents the relative probability of an observation (yield along y-axis) given a specific chill portion (along x-axis).

Estimated yield given the expected chill

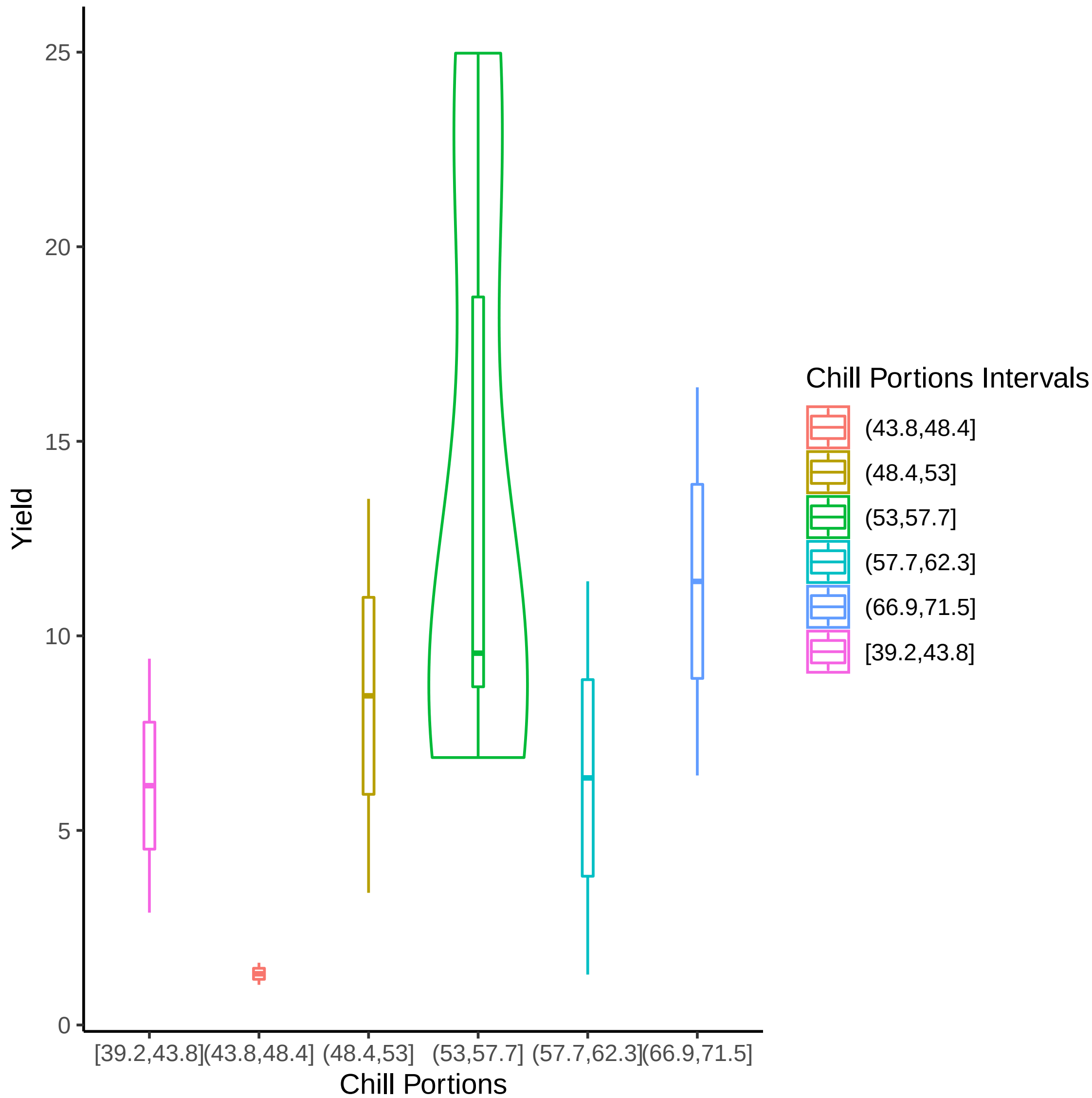
`chillkernelslice()` calculates the estimated yield given the expected chill, based on a slice of 'z' from the Kernel density calcualted with `chillkernel()`. The `expectedchill` parameter is set to 30.



`chillkernelslice()` plots the probabilities (shown along the y-axis) for the expected yield (shown along the x-axis). Since this is a cut through the density kernel `chillkernel()`, which integrates to 1, the probability values are relative, not absolute measures.

Chill portion intervals

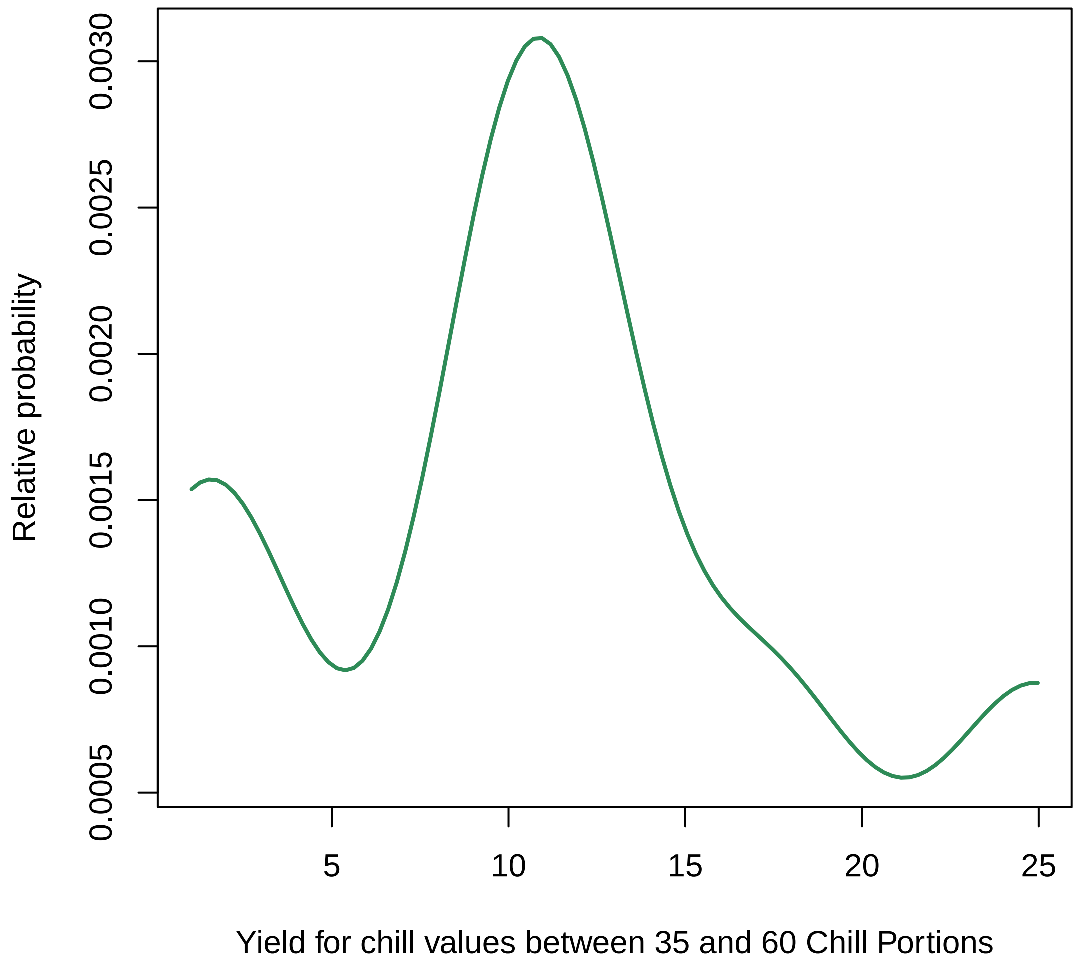
`chillviolin()` determines different possible chill portion intervals by calculating the optimal interval width for Chill Portions using the `IQR()` function in the `stats()` package, after the Freedman-Diaconis rule (IQR = interquartile range) (R Core Team, 2019). `Optimal interval width for our sample = 2 * interquartile range for our sample / (total number of observations for the interquartile range for our sample)^(1/3)`



`chillviolin()` shows violin plots of Chill Portions (x) and yield (y) with six different intervals of Chill Portions. Plot made with `ggplot2()` (Wickham et al., 2019).

Probability of yield given chill

`chillkernelslicerange()` The chill portion intervals, optimized interquartile ranges shown in `chillviolin()` can be used to select a range to slice from the density kernel `chillkernel()` as was doen for a single chill value in `chillkernelslice()`. Here we set the maximum chill to 35 and the minimum to 60.



`chillkernelslicerange()` plots the probabilities (shown along the y-axis) for the expected yield (shown along the x-axis). As with `chillkernelslice()` the probability values shown are relative, not absolute measures. They are the result of cuts through the density kernel `chillkernel()`, which integrates to 1.

Section

The functions here closely follow chillR (Luedeling, 2019) and decisionSupport (Luedeling et al., 2019).

References

Luedeling, E. 2019. ChillR: Statistical methods for phenology analysis in temperate fruit trees.

Luedeling, E., Goehring, L. and Schiffers, K. 2019. DecisionSupport: Quantitative support of decision making under uncertainty.

R Core Team. 2019. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ripley, B. 2019. MASS: Support functions and datasets for venables and ripley's mass.

Schiffers, K., Whitney, C., Fernandez, E. and Luedeling, E. 2018. PasitR: Calculates common functions for the pasit project.

Wickham, H. 2019. Plyr: Tools for splitting, applying and combining data.

Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C., Woo, K. and Yutani, H. 2019. Ggplot2: Create elegant data visualisations using the grammar of graphics.