

1 pep725: An R package for Pan-European Phenological 2 Data Analysis

3 **Barbara Templ**  ^{1*} and **Matthias Templ**  ^{2*} 

4 1 Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Switzerland  2 University

5 of Applied Sciences Northwestern Switzerland (FHNW), Switzerland  * Corresponding author *

6 These authors contributed equally.

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 



Editor: 

Submitted: 14 February 2026

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/))

9 Phenological observations are a key source of evidence for biological responses to climate variability and change (Parmesan, 2006; Piao et al., 2019; Thackeray et al., 2016). The Pan European Phenology database [PEP725; Templ et al. (2018); Templ et al. (2026)] is an open-access infrastructure for plant phenology data, unifying more than 13 million observations from over 30 countries, spanning the period from 1868 to the present and covering approximately 265 plant species and 46 phenophases. While these datasets offer exceptional scientific value, their analysis is challenged by uneven spatial coverage, heterogeneous data quality and temporal gaps. The **pep725** R package provides a coherent framework for the spatio-temporal analysis of phenological data from PEP725 and similar ground-based datasets. It supports standardized data import, systematic data quality diagnostics, computation of phenological normals and anomalies, estimation of spatial gradients, quantification of synchrony, climate–phenology sensitivity analysis, and both interactive and publication-ready visualization. In addition the package includes synthetic data generation tools to facilitate reproducible examples and teaching. Together, these components establish **pep725** as a specialized framework for quality diagnostic, spatio-temporal structure analyses, and visualization of ground-based phenological data.

25 Statement of need

26 Ground-based phenological datasets pose distinct analytical challenges: observation density varies across regions, long-term records are often incomplete, and data quality differs among contributors and time periods. Without explicit diagnostics and spatial characterization, such heterogeneity can bias spatial comparisons, trend analyses, and downstream model calibration.

30 While existing R packages primarily focus on process-based modeling, statistical estimation, or remote-sensing phenology, fewer tools address the diagnostic and spatio-temporal characterization of heterogeneous station-based datasets prior to modeling.

33 The **pep725** package addresses these needs by providing a dedicated framework (see Fig 34 1.) for characterizing station-based phenological datasets prior to modeling, separating data

35 diagnostics from downstream inference to enable more robust and reproducible analyses.

36 State of the field

37 Phenological analysis in R is supported by a diverse ecosystem of packages that address
38 complementary aspects of data modeling, prediction, and interpretation. Many tools focus
39 on process-based or climate-driven modeling of phenological events. For example, `chillR`
40 ([Luedeling, 2025](#)) provides a comprehensive framework for modeling chilling and forcing
41 requirements in temperate fruit trees, including bloom prediction and climate scenario analysis.
42 `DyMEP` ([Tschurr, 2025](#)) implements dynamic, weather-driven phenology models for major crops,
43 offering mechanistic prediction across environments. A general framework for phenology
44 modeling and forecasting is provided by `phenor` ([Hufkens et al., 2018](#)), which integrates
45 ground-based observations, remote-sensing products, and climate data to support predictive
46 and process-based applications. Crop-specific modeling approaches are further addressed
47 by `phenolocrop` ([Taniguchi, 2025](#)), which focuses on growth stage transitions and machine-
48 learning-based prediction.

49 Other packages emphasize statistical estimation of seasonal timing from irregular or sparse data.
50 `phenesse` ([Belitz et al., 2025](#)) applies quantile regression to estimate onset and end dates of
51 phenophases, enabling robust inference from opportunistic observations. The `phenology`
52 package ([Girondot, 2025](#)) provides parametric and Bayesian models for fitting seasonal
53 distributions of biological events. Auxiliary tools such as `pheno` ([Schaber & Lange, 2026](#))
54 and `spphpr` ([Shi et al., 2025](#)) support thermal-time-based phenological calculations and
55 preprocessing workflows.

56 A parallel line of development targets remote-sensing-based phenology, where land surface
57 phenology metrics are derived from vegetation index time series rather than ground observations.
58 Packages such as `npphen` ([Chavez et al., 2023](#)) and `phenex` ([Lange & Doktor, 2017](#)) provide
59 tools for detecting vegetation anomalies and analyzing spatial patterns in satellite-derived
60 phenology. The `sephora` package ([Tecuapetla-Gómez et al., 2024](#)) estimates phenological
61 transition dates from nearly periodic Earth Observation time series. In addition, `phenomap`
62 ([Zhang, 2020](#)) supports spatial interpolation and mapping of phenological observations.

63 Together, these tools provide extensive support for modeling, estimation, and remote-sensing
64 analysis of phenology. However, their primary focus lies on prediction, parameter estimation, or
65 satellite-derived metrics rather than on the systematic diagnostic and spatial characterization
66 of heterogeneous ground-based observation networks.

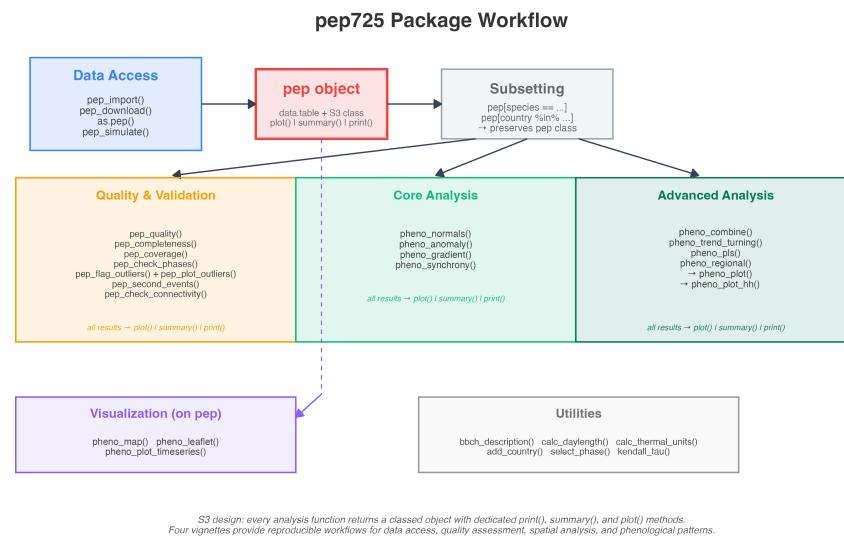
67 Software design

68 The design of `pep725` is guided by three principles: (i) explicit treatment of data quality and
69 uncertainty, (ii) transparent and reproducible spatio-temporal analysis workflows, and (iii)
70 modular integration with the broader R ecosystem. The package is built around an S3 class,
71 `pep`, implemented as an extended `data.table` with validated spatial and temporal attributes
72 (Fig 1.).

73 Data enter the workflow through `pep_import()`, `pep_download()`, `as.pep()`, `simulate_pep()`,
74 which create a unified `pep` object. Standard bracket notation (`pep[...]`) enables class-preserving
75 subsetting by species, region, phase, or time period.

76 The workflow separates quality and validation (e.g. completeness checks, coverage summaries,
77 outlier detection) from analytical procedures to ensure transparent intermediate diagnostics.
78 Analytical functionality is organized into core (e.g. normalization, anomalies, gradients,
79 synchrony) and advanced layers (e.g. trend modelling, combined metrics), with optional
80 robust estimators to reduce sensitivity to outliers and uneven sampling.

- 81 The package follows a consistent prefix-based naming convention. Functions operating directly
 82 on pep objects use the pep_ prefix, whereas more general phenological analysis and modelling
 83 routines use the pheno_ prefix.
- 84 A visualization layer provides maps and time-series plots directly from pep or result objects,
 85 while consistent S3 print(), summary(), and plot() methods support reproducible reporting.



86

87 Fig 1. pep725 package workflow and main functional components.

88 Example usage

- 89 The example below demonstrates a typical quality-first workflow.

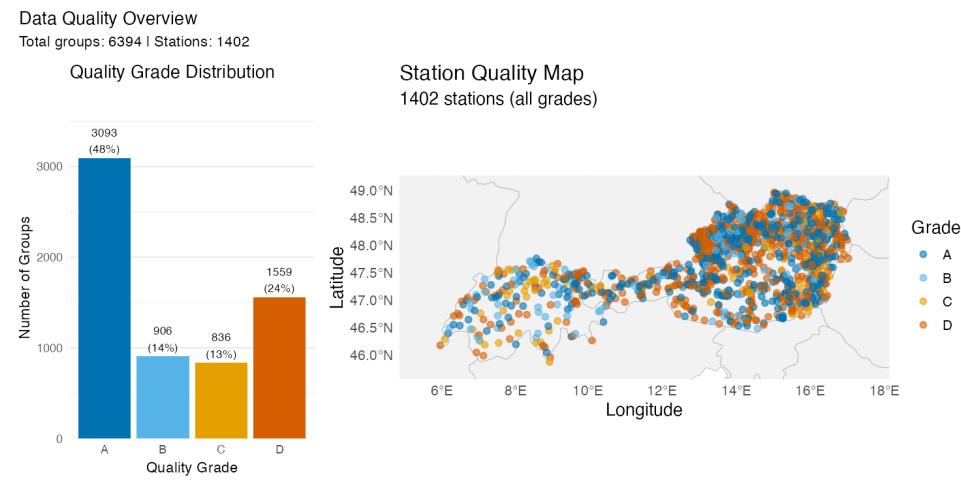
```

library(pep725)
pep <- pep_download()

# Alpine apple flowering
apple <- pep[species == "Malus domestica" &
             country %in% c("Switzerland", "Austria")]

# Grade each station's record quality (A-D)
quality <- pep_quality(apple, by = c("s_id", "phase_id"))
summary(quality)
plot(quality, which = "overview", pep = apple) # cf. Fig. 1
  
```

- 90 Another example provide an overview about the data quality for apple flowering stations in
 91 Austria and Switzerland.



92
93 Fig 2. Left: distribution of quality grades (A = best, D = poorest) across 6,394 station-phase
94 combinations. Right: geographic distribution of 1,402 stations colored by quality grade.

95 Research impact statement

96 The **pep725** package has been developed in close connection with the PEP725 database, which
97 has contributed to more than 115 peer-reviewed publications, including 17 in *Nature* and 2 in
98 *Science* (Templ et al., 2026). These studies address phenological questions that the package's
99 analytical functions directly support: climate sensitivity analyses such as Fu et al. (2015),
100 who demonstrated declining warming effects on spring leaf unfolding at 1,245 PEP725 sites
101 across Europe; and pan-European trend detection as in Menzel et al. (2020), who identified
102 attributable climate change fingerprints in 97,000 PEP725 time series spanning 1951–2018.

103 By implementing the methods described in Templ et al. (2026) as reproducible R workflows
104 with explicit quality diagnostics, **pep725** lowers the barrier to conducting such analyses on
105 large phenological observation networks. Four vignettes support both research applications
106 and teaching in phenology and environmental data analysis.

107 AI usage disclosure

108 AI-based tools were used as an editorial aid for language polishing and structural improvements
109 in parts of the documentation. All scientific concepts, methodological design, software
110 architecture, implementation, and validation were conceived, developed, and verified by the
111 authors.

112 Acknowledgements

113 We acknowledge the national phenological services and numerous observers contributing to
114 the PEP725 database. We also thank members of the phenology and climate-impact research
115 community for feedback during the development of this package. Dr. Barbara Templ was
116 supported by the Berner Nachwuchsförderung (BNF) program of the University of Bern for
117 developing the **pep725** R package. We thank Dr. Pierlugi Calanca at Agroscope (Switzerland)
118 for his inspirations and Dr. Yann Vitasse for hosting and supporting this work at the Swiss
119 Federal Institute for Forest, Snow and Landscape Research (WSL), which provided an excellent
120 research environment for conducting parts of the development and analysis.

121 References

- 122 Belitz, M., Campbell, C., & Li, D. (2025). *Phenesse: Estimate phenological metrics*
123 *using presence-only data* (Version 0.1.3) [Computer software]. <https://cran.r-project.org/package=phenesse>
- 125 Chavez, R. O., Estay, S. A., Lastra, J. A., Riquelme, C. G., Olea, M., Aquayo, J., & Decuyper, M.
126 (2023). *Npphen: Detect extreme vegetation anomalies from remotely sensed phenological*
127 *variability* (Version 2.0.1) [Computer software]. <https://doi.org/10.3390/rs15010073>
- 128 Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel,
129 A., Peñuelas, J., Song, Y., Vitassee, Y., Zeng, Z., & Janssens, I. A. (2015). Declining
130 global warming effects on the phenology of spring leaf unfolding. *Nature*, 526, 104–107.
131 <https://doi.org/10.1038/nature15402>
- 132 Girondot, M. (2025). *Phenology: Tools to manage a parametric function that describes*
133 *phenology* (Version 2025.11.12) [Computer software]. <https://doi.org/10.3354/esr00292>
- 134 Hufkens, K., Basler, D., Milliman, T., Melaas, E. K., & Richardson, A. D. (2018). *Phenor:*
135 *An integrated phenology modelling framework* (Version 1.3.1) [Computer software]. <https://doi.org/10.1111/2041-210X.12970>
- 137 Lange, M., & Doktor, D. (2017). *Phenex: Auxiliary functions for phenological data analysis*
138 (Version 1.4-5) [Computer software]. <https://cran.r-project.org/package=phenex>
- 139 Luedeling, E. (2025). *chillR: Statistical methods for phenology analysis in temperate fruit trees*
140 (Version 0.77) [Computer software]. <https://doi.org/10.1016/j.scienta.2012.07.011>
- 141 Menzel, A., Yuan, Y., Matiu, M., Sparks, T., Scheifinger, H., Gehrig, R., & Estrella, N. (2020).
142 Climate change fingerprints in recent european plant phenology. *Global Change Biology*,
143 26(4), 2599–2612. <https://doi.org/10.1111/gcb.15000>
- 144 Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual*
145 *Review of Ecology, Evolution, and Systematics*, 37, 637–669. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110100>
- 147 Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu,
148 X. (2019). Plant phenology and global climate change: Current progresses and challenges.
149 *Global Change Biology*, 25(6), 1922–1940. <https://doi.org/10.1111/gcb.14619>
- 150 Schaber, J., & Lange, M. (2026). *Pheno: Auxiliary functions for phenological data analysis*
151 (Version 1.7-1) [Computer software]. <https://cran.r-project.org/package=pheno>
- 152 Shi, P., Chen, Z., Reddy, G. V. P., Hui, C., Huang, J., & Xiao, M. (2025). *Spphpr: Spring*
153 *plant phenology prediction* (Version 1.1.5) [Computer software]. <https://doi.org/10.1016/j.agriformat.2017.04.001>
- 155 Taniguchi, Y. (2025). *Phenolocrop: Crop phenology modelling and parameterization* (Version
156 0.0.4) [Computer software]. <https://cran.r-project.org/package=phenolocrop>
- 157 Tecuapetla-Gómez, I., Galicia-Gómez, F., & Rosales-Marticorena, F. (2024). *Sephora:*
158 *Statistical estimation of phenological parameters* (Version 0.1.31) [Computer software].
159 <https://cran.r-project.org/package=sephora>
- 160 Templ, B., Koch, E., Bolmgren, K., Ungersböck, M., Paul, A., Scheifinger, H., Rutishauser,
161 T., Busto, M., Chmielewski, F.-M., Dobrovolný, P., Gregor, M., Hajkova, L., Hajkova,
162 V., Huep, J., Kaspar, F., Pietragalla, B., Romero-Fresneda, R., Tolvanen, A., Vučetić,
163 V., ... Menzel, A. (2018). Pan european phenological database (PEP725): A single point
164 of access for european data. *International Journal of Biometeorology*, 62(6), 1109–1113.
165 <https://doi.org/10.1007/s00484-018-1512-8>
- 166 Templ, B., Scheifinger, H., Ostovary, I., Ungersbock, M., & Ressl, H. (2026). PEP725:

- 167 15 years of driving european and global phenology science. *New Phytologist*. <https://doi.org/10.1111/nph.70869>
- 168
- 169 Thackeray, S. J., Henrys, P. A., Hemming, D., Bell, J. R., Botham, M. S., Burthe, S., Helaouet, P., Johns, D. G., Jones, I. D., Leech, D. I., Mackay, E. B., Massimino, D., Atkinson, S., Bacon, P. J., Brereton, T. M., Carvalho, L., Clutton-Brock, T. H., Duck, C., Edwards, M., ... Wanless, S. (2016). Phenological sensitivity to climate across taxa and trophic levels. *Nature*, 535, 241–245. <https://doi.org/10.1038/nature18608>
- 170
- 171
- 172
- 173
- 174 Tschurr, F. (2025). *DyMEP: Dynamic multi-environment phenology model* (Version 0.1.2) [Computer software]. <https://cran.r-project.org/package=DyMEP>
- 175
- 176 Zhang, X. (2020). *Phenomap: Remote sensing based phenological mapping* (Version 2.0.1) [Computer software]. <https://cran.r-project.org/package=phenomap>
- 177

DRAFT