Temporal Systems and Forecasting
Volume 1 · Number 1 · 2026

A Small Wind-Forecast Advantage Across a Later Year: Quarters and Reference Choices

William Murphy[1] · Lucas Walker[2] · Evan Anderson[3]

[1] Austria · [2] Belgium · [3] Italy

Original publication: November 4, 2026 | Article 669

Abstract

Adding neighboring turbine readings lowered average next-interval wind-speed error in every quarter of a later-year holdout, but the advantage was small and uneven. Models fitted on 2014 records were evaluated on 51,384 ten-minute origins in 2015, summarized into 361 available dates. The neighboring-input forecast improved equal-date absolute error by 0.010246 m/s relative to an own-turbine linear forecast. Quarterly gains ranged from 0.002558 to 0.015765 m/s. Differences among quarters were uncertain under that reference; some became distinguishable when persistence supplied the reference instead. The result supports a limited example of chronological forecasting benefit while showing that temporal interpretation depends on the comparison forecast and the uncertainty attached to it.

Frozen linear forecasts on public La Haute Borne turbine records; a chronological comparison of wind-speed errors rather than power-generation performance.

A Small Wind-Forecast Advantage Across a Later Year: Quarters and Reference Choices

Frozen linear forecasts on public La Haute Borne turbine records; a chronological comparison of wind-speed errors rather than power-generation performance.

Abstract

Adding neighboring turbine readings lowered average next-interval wind-speed error in every quarter of a later-year holdout, but the advantage was small and uneven. Models fitted on 2014 records were evaluated on 51,384 ten-minute origins in 2015, summarized into 361 available dates. The neighboring-input forecast improved equal-date absolute error by 0.010246 m/s relative to an own-turbine linear forecast. Quarterly gains ranged from 0.002558 to 0.015765 m/s. Differences among quarters were uncertain under that reference; some became distinguishable when persistence supplied the reference instead. The result supports a limited example of chronological forecasting benefit while showing that temporal interpretation depends on the comparison forecast and the uncertainty attached to it.

1. Transport into a new year is more than a single error average

A fitted wind forecast may have a positive overall advantage over a reference model while that advantage varies substantially through the year. An annual average can conceal a period in which additional inputs help little. It can also obscure a second comparison problem: conclusions about stability depend on which reference forecast supplies the subtraction. A model may look consistently better than a weak alternative and less consistent against a simple persistence forecast.

We ask whether a fixed model using contemporary readings from neighboring turbines retains its absolute-error advantage across quarters of the next calendar year. The primary comparison is an own-turbine linear forecast estimated on the same training period. Persistence provides a second, untuned reference. The directional proposition is that adding neighbors yields negative daily loss differences throughout the later year. A separate proposition concerns whether quarter-specific differences depart from the first-quarter contrast. These questions require distinct interpretation even when both are summarized in a regression.

The wind-prediction context of Yan and Kuruoglu, 2023 motivates examining robustness under later observations. Our model contains ordinary linear terms and is fitted only once. It does not implement graph-sign diffusion, inject impulsive disturbances, or assess that method's performance. The present contribution is a transparent temporal comparison using observed turbine records. It supplies no evidence that an adaptive graph algorithm would have the same gains, weaknesses, or response to a changing environment.

2. The forecasting task and what the source can support

Operational forecasting issues discussed by Pinson, 2013 and the methods review in Foley et al., 2012 motivate taking horizon, information availability, and evaluation period seriously. We predict wind speed at a single turbine ten minutes after the current timestamp. This is not electrical power, a probabilistic wind forecast, or an energy-market decision. A difference measured in meters per second cannot be converted into revenue or power-output improvements without additional turbine and operational information.

The spatial perspective in Haslett and Raftery, 1989 helps motivate neighboring observations as candidate inputs, but our fitted model does not estimate a physical flow field or long-memory process. The four contemporaneous readings form a small multivariate information set. Correlation can make neighboring measurements useful for prediction without identifying an upstream causal influence. We do not infer aerodynamic mechanisms from fitted coefficients or regard turbines as independently randomized experimental units.

Evaluation guidance in Tashman, 2000 makes chronological separation central to this design. Every model coefficient is estimated from targets before 2015; none is refitted during the holdout year. Cross-validation concerns in Bergmeir et al., 2018 and dependence considerations in Roberts et al., 2017 explain why an arbitrary random division of adjacent ten-minute observations would answer a different question. We retain the later year as a sequential test period and report variation within it, rather than presenting a random-split score as evidence of transport.

3. Public records, conflicting keys, and the forecast origin

The observations are the La Haute Borne 2014-2015 SCADA wind-speed records captured in ENGIE / NREL OpenOA contributors, 2026, supplied within the public example archive associated with Perr-Sauer et al., 2021. The original wind table has 420,480 rows across four turbines. We parse timestamps in UTC and check turbine-time uniqueness before constructing a wide table. Forty-eight keys have conflicting duplicate records, corresponding to ninety-six source rows. We exclude all rows belonging to those ambiguous keys, avoiding an undocumented first-value or last-value choice.

The target turbine is R80711. Its current wind speed is the own-turbine predictor; the other three turbine readings are candidate neighboring inputs. We retain only origins at which all four current values and the next target value are observed, and require the target timestamp to be exactly ten minutes later. Rows separated by an interruption in the source time index cannot become ordinary next-interval examples merely because they are adjacent in a sorted file. No missing wind value is interpolated.

Training retains 52,321 origins whose target timestamps occur before 1 January 2015. Testing retains 51,384 origins dated during 2015, including only targets with the required next-interval match. This target-based training boundary prevents an origin just before midnight from importing the following year's target into estimation. All three forecasts are evaluated on these identical held-out rows. An own-only forecast could operate on a broader availability set, but using that broader set would mix information-set effects with a change in evaluation observations.

4. Three forecasts with two estimated coefficient sets

The own-only forecast is ordinary least squares with an intercept and current own wind speed. The neighboring-input forecast adds the other three contemporary wind-speed readings, again with an intercept. Both are estimated solely on the retained training origins. Persistence predicts the future wind speed by repeating the current own reading, with no estimated slope or intercept. There are no wind-direction, temperature, calendar, hour, or lagged-error predictors. The numerical design matrices of the two fitted models have full column rank.

The own-only intercept is approximately 0.175970 m/s and its slope is 0.968328. The neighboring-input intercept is 0.161803; its own-speed slope is 0.749830, with neighboring slopes about 0.075367, 0.101622, and 0.059815. These coefficients are provided to describe the actual frozen forecasts. Their sizes do not establish physical transport weights or a causal contribution of each turbine. Correlated current readings and source-specific calibration can affect how ordinary least squares distributes coefficients.

Predictive comparison principles in Diebold and Mariano, 1995 and West, 1996 distinguish evaluating losses from interpreting a forecasting coefficient as an explanation. The later discussion in Diebold, 2015 also motivates caution about treating a named accuracy test as a universal solution to forecast comparison. We report direct paired loss differences and a declared covariance calculation. We do not call this a repeated-origin model-selection experiment or infer that one fixed fitted model dominates all plausible forecasting methods.

5. Daily loss differences preserve the paired comparison

For each held-out origin we calculate absolute forecast error in meters per second for all three forecasts. The neighboring forecast's error minus the own-only error defines the primary contrast. Its error minus persistence defines the secondary contrast. Negative values favor the neighboring forecast. Each contrast is first averaged within a UTC date, giving 361 retained dates. Days with incomplete coverage are included, and every retained date then receives equal weight in the analysis.

The accuracy-measure discussion in Hyndman and Koehler, 2006 matters because equal-date mean absolute error differs from weighting every ten-minute origin equally. We select the former to keep a heavily observed day from automatically dominating a partly observed one. This choice does not recover error for the four dates with no eligible origins, nor does it remove bias if observation availability is related to difficult weather. Missing-data reasoning in Little and Rubin, 2002 makes that unobserved portion a limit rather than an invitation to fill the error series with zeros.

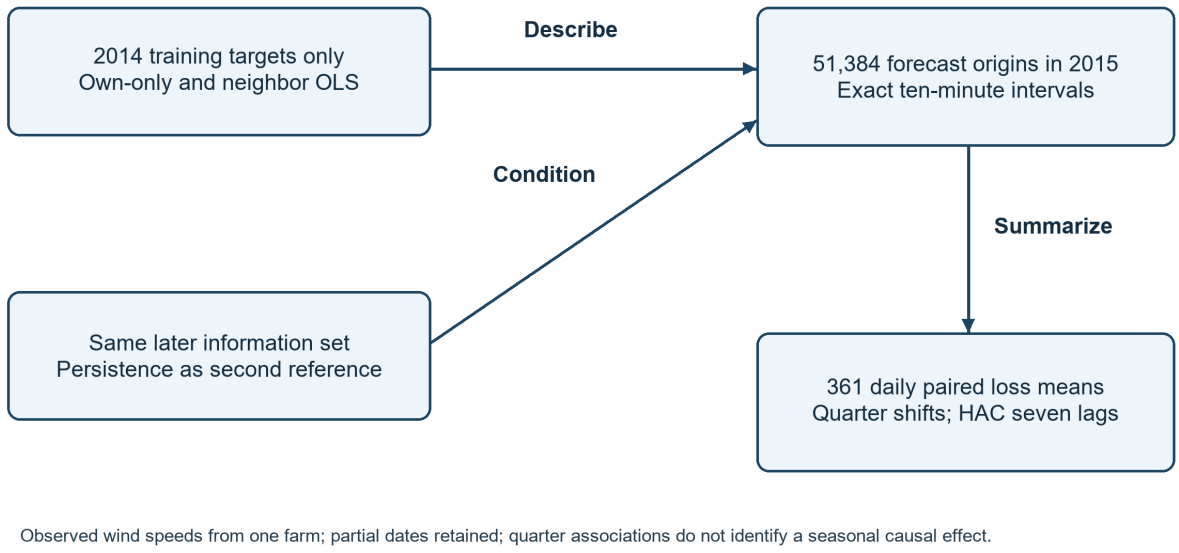


Figure 1. Two 2014-fitted linear forecasts and a persistence rule predict identical 2015 origins; paired errors become daily summaries before quarter comparisons.

The diagram ends with held-out losses. Scoring principles in Gneiting and Raftery, 2007 and calibration concerns in Guo et al., 2017 would be relevant to a probabilistic forecast extension, but no predictive distributions or coverage estimates are produced here. The measurement framework of Carroll et al., 2006 also reminds us that released anemometer readings are observed inputs and targets. Their physical calibration is not independently verified by a low forecast error. We preserve the source units rather than claim access to error-free atmospheric wind.

6. A quarter regression with serially robust uncertainty

For each daily contrast we fit ordinary least squares on an intercept and indicators for quarters two, three, and four. Quarter one is the reference. The intercept gives the first-quarter mean loss difference; each quarter coefficient gives a departure from that first-quarter difference. Both regressions use all 361 daily rows. Their R-squared values, 0.049143 and 0.074948, describe variation in daily error contrasts, not the accuracy or explained variance of the original wind-speed forecasts.

Covariance is heteroskedasticity- and autocorrelation-consistent with seven available-date lags, a finite-sample correction, and a t reference using 357 residual degrees of freedom. The approach draws on Newey and West, 1987. Seven retained lags are not exactly seven calendar days when observations are absent. Heteroskedasticity-consistent reasoning in White, 1980 alone would omit serial dependence, but the chosen HAC specification also remains a modeling choice rather than a guarantee that all seasonal or long-range dependence has been captured.

Table 1. Later-quarter shifts in neighboring-forecast daily absolute-error differences, with quarter one as the reference.

Model	Term	Estimate (SE)	95% CI	N
M1	Quarter 1 mean contrast	-0.0076 (0.0050)	[-0.0175, 0.0023]	361
M1	Quarter 2 minus quarter 1	-0.0073 (0.0051)	[-0.0173, 0.0027]	361
M1	Quarter 3 minus quarter 1	-0.0081 (0.0053)	[-0.0186, 0.0023]	361
M1	Quarter 4 minus quarter 1	0.0051 (0.0059)	[-0.0065, 0.0166]	361
M1: fit	R-squared = 0.0491	-	-	361
M2	Quarter 1 mean contrast	-0.0058 (0.0052)	[-0.0160, 0.0044]	361
M2	Quarter 2 minus quarter 1	-0.0116 (0.0052)	[-0.0218, -0.0014]	361
M2	Quarter 3 minus quarter 1	-0.0137 (0.0055)	[-0.0245, -0.0029]	361
M2	Quarter 4 minus quarter 1	0.0022 (0.0059)	[-0.0094, 0.0139]	361
M2: fit	R-squared = 0.0749	-	-	361

Model labels: M1 = Neighbor minus own-only loss by later quarter; M2 = Neighbor minus persistence loss by later quarter.

We do not fit a multilevel trajectory model such as those motivated by Laird and Ware, 1982, because the evaluated loss series concerns one target turbine in one wind farm. Nor do four quarters supply four independent field replications. The quarter indicators summarize a dated sequence from this resource. The confidence intervals cannot establish geographic transport to a new wind farm or provide the uncertainty of repeatedly acquiring and refitting an entirely new forecasting dataset.

7. The pooled gain is small, and its point estimates vary

Across equal-weight retained dates, the own-only forecast has mean absolute error 0.433715403 m/s, the neighboring forecast 0.423469520 m/s, and persistence 0.435087604 m/s. The neighboring forecast's differences are therefore -0.010245883 m/s against own-only and -0.011618084 m/s against persistence. These pooled values are descriptive means. The primary improvement is about one hundredth of a meter per second, so its practical importance cannot be assumed from the number of evaluated ten-minute origins.

Table 2. Quarter-specific means on the same daily evaluation series, in meters per second.

Quarter	Retained dates	Own-only MAE	Neighbor MAE	Persistence MAE	Neighbor minus own-only
1	86	0.433431	0.425802	0.431583	-0.007629
2	91	0.434908	0.419996	0.437356	-0.014911
3	92	0.474072	0.458307	0.477776	-0.015765
4	92	0.392446	0.389887	0.393432	-0.002558

All four primary point estimates favor adding neighbors. The smallest advantage is in the final quarter, while the largest is in the third. Quarter one's estimated contrast is -0.007629 with standard error 0.005044 and interval -0.017548 to 0.002290. Its interval contains zero. The changes from quarter one are -0.007283, -0.008136, and 0.005070 for quarters two, three, and four respectively; all three corresponding intervals also contain zero. Descriptive variation is therefore clearer than precise evidence of a quarter departure under the primary reference.

8. Changing the reference alters the temporal contrast

Against persistence, the first-quarter mean difference is -0.005780, with interval -0.015993 to 0.004432. The second-quarter departure from that difference is -0.011579, with interval -0.021788 to -0.001370. The third-quarter departure is -0.013689, with interval -0.024495 to -0.002882. The fourth-quarter departure is 0.002235, with interval -0.009426 to 0.013897. Two quarter departures thus exclude zero under the secondary reference despite the uncertain departures in the own-only comparison.

These coefficients compare quarter-specific advantages with the first-quarter advantage. An interval for a departure is not automatically an interval for that quarter's total mean gain. The result should not be reported as a direct significance test of every quarter's neighboring forecast against zero loss difference. Table 2 supplies primary mean levels, while Table 1 supplies the actual regression quantities and their uncertainty. Keeping these objects separate prevents a common misreading of indicator regressions.

No evidence here identifies why persistence changes relative to a fitted own-only model over the year. Their errors can respond differently to the distribution of speeds, mean reversion, measurement availability, or other changes. We have not decomposed those mechanisms. Guidance from Wasserstein and Lazar, 2016 favors reporting the two reference choices and effect magnitudes instead of declaring a stable or unstable forecast solely from which individual interval excludes zero. Multiplicity considerations in Benjamini and Hochberg, 1995 further limit a strong claim based on a collection of quarter coefficients.

9. A calendar pattern cannot identify a seasonal mechanism

The causal-design perspectives in Greenland et al., 1999 and Stuart, 2010 help delimit this interpretation. Quarter is a calendar label, not a randomized exposure. It can proxy for weather regimes, instrument changes, availability, or unmeasured operational differences. Adding a quarter indicator does not isolate a causal effect of season. We make no claim that moving an observation into another quarter would change the model's advantage by the estimated coefficient.

Model-selection bias discussed in Varma and Simon, 2006 also constrains reuse of this holdout. If these quarterly outcomes were used to redesign predictors, choose a lag, or schedule model updates, they would become development evidence for that new policy. A subsequent independent test would then be needed. Influence reasoning associated with Cook, 1977 could motivate checking whether unusual days drive particular contrasts; no high-error date is removed from this report. We retain partial dates and acknowledge that their inclusion defines the current equal-date target.

Deployment considerations in Sculley et al., 2015 are relevant because the calculation assumes all contemporary neighboring readings arrive before a forecast is issued. Real communication latency or sensor failure could change that information set. Neither has been measured. The reliability concerns raised in Ioannidis, 2005 likewise favor replication on another period or farm before making a general operational recommendation. One public wind farm and one later year offer a useful inspectable case, rather than an independently replicated account of forecasting robustness.

10. What should accompany a reported annual advantage

The evidence supports a small average reduction in absolute wind-speed error from adding neighboring readings to this frozen linear predictor. It also shows why an annual advantage needs a dated breakdown and a clearly named reference. Quarterly point gains vary appreciably, while uncertainty and the choice of own-only versus persistence change what can be said about those differences. No graph-algorithm ranking, economic benefit, or causal explanation is needed to state that result precisely.

The documentation aims of Gebru et al., 2021 are served by preserving duplicate-key exclusions, the exact ten-minute matching rule, target-based training cutoff, shared availability sample, and partial-date weighting. Following Wilkinson et al., 2016, the archived records, held-out predictions, daily paired losses, and quarter summaries remain linked to the executable analysis. These materials make the numerical comparisons reproducible. Broader claims about adaptation, geographic generalization, or operational value require additional evidence beyond this completed chronological reanalysis.

References

1. Hyndman, Rob J. and Anne B. Koehler. “Another look at measures of forecast accuracy.” *International Journal of Forecasting*, vol. 22, no. 4, 2006, pp. 679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001.
2. Newey, Whitney K., and Kenneth D. West. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica*, vol. 55, no. 3, 1987, pp. 703-708. https://doi.org/10.2307/1913610.
3. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. “Measurement Error in Nonlinear Models.” 2006. https://doi.org/10.1201/9781420010138.
4. Benjamini, Yoav and Yosef Hochberg. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 57, no. 1, 1995, pp. 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
5. Pinson, Pierre. (2013). Wind Energy: Forecasting Challenges for Its Operational Management. Statistical Science. https://doi.org/10.1214/13-sts445.

6. ENGIE / NREL OpenOA contributors, . (2026). La Haute Borne 2014-2015 SCADA and variable dictionary: captured OpenOA example archive. Local snapshot captured 7 November 2026; observed data years 2014-2015. https://github.com/NREL/OpenOA/blob/main/examples/data/la_haute_borne.zip.

7. Diebold, Francis X.. (2015). Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. Journal of Business & Economic Statistics. https://doi.org/10.1080/07350015.2014.983236.

8. White, Halbert. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, vol. 48, no. 4, 1980, pp. 817-838. https://doi.org/10.2307/1912934.

9. Little, Roderick J. A., and Donald B. Rubin. *Statistical Analysis with Missing Data*. Wiley, 2002. https://doi.org/10.1002/9781119013563.

10. Gneiting, Tilmann; Raftery, Adrian E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. Journal of the American Statistical Association. https://doi.org/10.1198/016214506000001437.

11. Wasserstein, Ronald L.; Lazar, Nicole A.. (2016). The ASA Statement on p -Values: Context, Process, and Purpose. The American Statistician. https://doi.org/10.1080/00031305.2016.1154108.

12. Stuart, Elizabeth A. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science*, vol. 25, no. 1, 2010, pp. 1-21. https://doi.org/10.1214/09-sts313.

13. Cook, R. Dennis. (1977). Detection of Influential Observation in Linear Regression. Technometrics. https://doi.org/10.1080/00401706.1977.10489493.

14. Gebru, Timnit, et al. "Datasheets for datasets" *Communications of the ACM*, 2021. https://doi.org/10.1145/3458723.

15. Roberts, David R., et al. “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure.” *Ecography*, vol. 40, no. 8, 2017, pp. 913-929. https://doi.org/10.1111/ecog.02881.

16. Varma, Sudhir and Richard Simon. “Bias in error estimation when using cross-validation for model selection.” *BMC Bioinformatics*, vol. 7, no. 1, 2006. https://doi.org/10.1186/1471-2105-7-91.

17. Laird, Nan M. and James H. Ware. “Random-Effects Models for Longitudinal Data.” *Biometrics*, vol. 38, no. 4, 1982, pp. 963. https://doi.org/10.2307/2529876.

18. West, Kenneth D.. (1996). Asymptotic Inference about Predictive Ability. Econometrica. https://doi.org/10.2307/2171956.

19. Perr-Sauer, Jordan; Optis, Mike; Fields, Jason; Bodini, Nicola; Lee, Joseph; Todd, Austin; Simley, Eric; Hammond, Robert; Phillips, Caleb; Lunacek, Monte; Kemper, Travis; Williams, Lindy; Craig, Anna; Agarwal, Nathan; Sheng, Shawn; Meissner, John. (2021). OpenOA: An Open-Source Codebase For Operational Analysis of Wind Farms. Journal of Open Source Software. https://doi.org/10.21105/joss.02171.

20. Ioannidis, John P. A.. (2005). Why Most Published Research Findings Are False. PLoS Medicine. https://doi.org/10.1371/journal.pmed.0020124.

21. Haslett, John; Raftery, Adrian E.. (1989). Space-Time Modelling with Long-Memory Dependence: Assessing Ireland's Wind Power Resource. Applied Statistics. https://doi.org/10.2307/2347679.

22. Sculley, D., et al. "Hidden Technical Debt in Machine Learning Systems." *Advances in Neural Information Processing Systems*, vol. 28, 2015. https://papers.nips.cc/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html

23. Tashman, Leonard J.. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. International Journal of Forecasting. https://doi.org/10.1016/s0169-2070(00)00065-0.

24. Yan, Yi; Kuruoglu, Ercan E.. (2023). Fast and Robust Wind Speed Prediction Under Impulsive Noise via Adaptive Graph-Sign Diffusion. 2023 IEEE Conference on Artificial Intelligence (CAI), 302-305. https://doi.org/10.1109/CAI54212.2023.00135.

25. Greenland, Sander, Judea Pearl, and James M. Robins. “Causal Diagrams for Epidemiologic Research.” *Epidemiology*, vol. 10, no. 1, 1999, pp. 37-48. https://doi.org/10.1097/00001648-199901000-00008.

26. Guo, Chuan, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. “On Calibration of Modern Neural Networks.” *Proceedings of the 34th International Conference on Machine Learning*, PMLR, vol. 70, 2017, pp. 1321-1330. https://proceedings.mlr.press/v70/guo17a.html.

27. Foley, Aoife M.; Leahy, Paul G.; Marvuglia, Antonino; McKeogh, Eamon J.. (2012). Current methods and advances in forecasting of wind power generation. Renewable Energy. https://doi.org/10.1016/j.renene.2011.05.033.

28. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; 't Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.

29. Bergmeir, Christoph; Hyndman, Rob J.; Koo, Bonsoo. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. Computational Statistics & Data Analysis. https://doi.org/10.1016/j.csda.2017.11.003.

30. Diebold, Francis X.; Mariano, Roberto S.. (1995). Comparing Predictive Accuracy. Journal of Business & Economic Statistics. https://doi.org/10.1080/07350015.1995.10524599.