



Configural Analysis of Oscillating Progression

Alexander von Eye¹, Wolfgang Wiedermann², and Stefan von Weber³

¹ Michigan State University, USA

² University of Missouri, Columbia, USA

³ University of Furtwangen, Germany

Corresponding author:

Alexander von Eye, 190 Allée du Nouveau Monde, 34000 Montpellier, France

E-mail: voneye@msu.edu

To cite this article: von Eye, A., Wiedermann, W., & von Weber, S. (2021). Configural analysis of oscillating progression. *Journal for Person-Oriented Research*, 7(2), 14-21. <https://doi.org/10.17505/jpor.2020.23448>

Abstract

Oscillating series of scores can be approximated with locally optimized smoothing functions. In this article, we describe how such series can be approximated with locally estimated (loess) smoothing, and how Configural Frequency Analysis (CFA) can be used to evaluate and interpret results. Loess functions are often hard to describe because they cannot be represented by just one function that has interpretable parameters. In this article, we suggest that specification of the CFA base model be based on the width of the window that is used for local curve optimization, the weight given to data points in the neighborhood of the approximated one, and by the function that is used to locally approximate observed data. CFA types indicate that more cases were found than expected from the local optimization model. CFA antitypes indicate that fewer cases were found. In a real-world data example, the development of Covid-19 diagnoses in France is analyzed for the beginning period of the pandemic.

Keywords: Configural Frequency Analysis, base model, loess smoothing, time series, local optimization, Covid-19

In many contexts of real-world data analysis, series of observations do not follow a simple straight-lined or smoothly curved progression. Instead, they suggest ups and downs and oscillations of various sorts. Examples of such progression include brain waves, stock market development, heart beats, and traffic load on bridges. Another empirical example (which will be analyzed in detail later in this article) can be found in the progression of positive Covid-19 diagnoses. Figure 1 displays this development over 83 days, beginning with the first diagnosis in France on January 24, 2020 (Santé Publique, 2020).

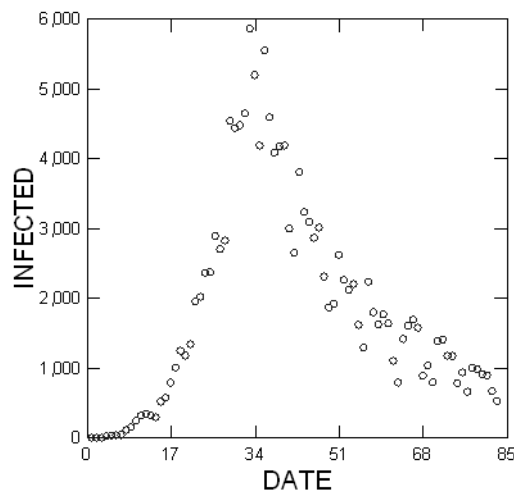
Figure 1 shows that there clearly is a strong, accelerated upwards trend for the first 36 days that is followed by a downward trend. In addition, this trend is not smooth, in particular after the number of positive diagnoses has reached its first peak, at day 36.

In this article, we (1) propose approximating this kind of series of observations with loess smoothing. We (2) also propose using Configural Frequency Analysis (CFA) to evaluate the smoothed curve, and to identify segments in which an otherwise well-functioning selection of smoothing para-

meters fails to properly describe the series.

Figure 1

Number of positive Covid-19 diagnoses in France in the 83 days after the first positive diagnosis



This article is structured as follows. First, we provide an overview of loess smoothing. Second, we discuss evaluation and interpretation of loess-smoothed curves with CFA. Third, we analyze in detail the data presented in Figure 1.

Local Weighted Regression (Loess)

Originally (Cleveland, 1979; Cleveland, & Devlin, 1988; see also Wilkinson, Blank, & Gruber, 1996), the smoother to be discussed in this article was called *lowess*, brief for **l**ocally **w**eighted scatterplot smoother. The method is also known as Savitzky-Golay filter (Savitzky & Golay, 1964). Now, the very same method is called *locally weighted regression smoothing* or *locally weighted scatterplot smoothing*, and is abbreviated with *loess*. To introduce this smoother, we follow Dagum and Luati (2003; see also Dagum & Luati, 2000, 2001) and first note that a series of scores can have a *global* and a *local representation*. When the global representation is used, one model is used for the entire series of scores. For example, one straight regression line or polynomial is used to describe the entire series (for the use of polynomials in CFA, see, e.g., von Eye & Wiedermann, 2021). When local representation is used, the series of scores is subdivided in sectors (the windows), and regression lines, polynomials, or other functions are used to describe the scores in each sector. Typically, the function that is used is the same for all sectors.

In standard applications of the General Linear Model (GLM; e.g., regression or ANOVA applications), weights can be used to reduce the error of a model. When ordinary least squares is used to estimate GLM parameters, the parameter vector, β , is

$$\hat{\beta} = (X'WX)^{-1}X'WY,$$

where X is the design matrix, Y denotes the observed scores, and W is the weight matrix. This equation holds for both global and local representations of series of scores.

The design matrix contains the scores of the predictors and/or the coefficients of the polynomials used to model the series of scores. For example, a regression in which a square curvature is fitted, could be

$$\hat{y} = b_0 + b_1x + b_2x^2,$$

where the x are the predictor scores, and the b are the regression parameters. In the context of modeling series of scores, the x often are the observation points in time.

When a global representation of a series is intended, this type of model is applied to the entire series of scores. When regression weights are used, these are often the scores of third variables, observed at the same points in time. When, however, loess smoothing is used, two elements of this approach differ. First, the regression model is applied only to the scores that lie within a predefined window. The size of the window is a priori determined and can be measured in units of number of observation points. Second, the weights are also measured in units of observation points. More specifically, the observation points close to the data point to be

approximated receive greater weights than the observation points farther away.

In more technical terms, let the observation point to be modeled be t_j , where $j \in J$ numbers the observation points, and let t_k be the observation points that are time adjacent and within the window. Then, one weight w_k of t_k is, for t_j ,

$$w_k(t_j) = \frac{1 - |t_k - t_j|}{\Delta(t_j)},$$

where $\Delta(t_j) = \max|t_j - t_k|$. Evidently, the weight depends on the size of the window and, thus, the distance between observation points within this window.

Alternatively, and traditionally, the weight function for loess smoothing is the *tri-cube* weight function,

$$w_k = \left(1 - |t_k - t_j|^3\right)^3.$$

In loess smoothing, each window is of equal width. That is, each window contains the same number of points. The smoothing parameter is the portion of observation points within a window. The wider the window, the smoother the fitted curve.

When the absolute value of the distance $t_k - t_j$ is scaled to range from 0 to 1, the tri-cube weight function assumes the form depicted in Figure 2.

Figure 2

Tri-cube weight function for $0 \leq |d| \leq 1$

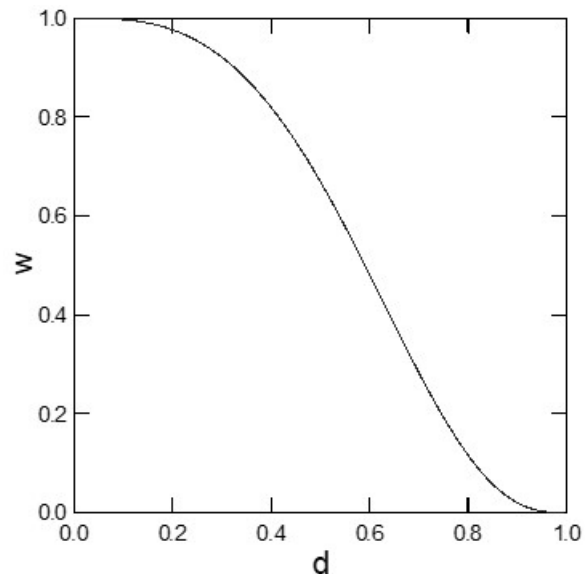


Figure 2 indicates that the weight decreases as the distance increases.

Within each window, the same polynomial regression is performed. The degree of the polynomial is determined a priori and is bounded by the width of the window. To prevent hitting the points exactly – which would counter the idea of smoothing – the polynomial should be of a degree of ‘number of observation points within the window – 2.’ Dagum and Luati (2003) state that low degree polynomials, e.g., linear

or quadratic polynomials, are usually the best choice. Higher degree polynomials are suitable in particular when the smoothing is expected to reflect at least some of the observed oscillations, and when the window includes many data points.

The width of the window is also determined a priori. Selecting narrow windows results in a curve that oscillates with the observed scores. Selecting a wide window will result in a very smooth curve with no oscillations.

Given the width of the window, the weights, the degree of the polynomial, and the magnitude of the smoothing parameter, the weighted least squares regression parameter for window j is estimated as

$$\hat{\beta}_j = (X_j'W_jX_j)^{-1}X_j'W_jY_j.$$

That is, one weighted polynomial regression is estimated for the series of scores in each window.

CFA and Loess Smoothing

Configural frequency analysis (CFA; Lienert, 1968; for technical elements and detail on CFA and recent developments, see von Eye et al., 2010, von Eye & Wiedermann, 2021, and Wiedermann et al., 2021) allows one to identify strong deviations from model-based expected frequencies. When, locally, significantly more cases are observed than expected based on the model, a CFA *type* has been identified. When fewer cases are observed, a CFA *antitype* has been identified.

With only very few exceptions, CFA is applied to the analysis of cross-classifications of two or more variables. The exceptions are all analyses of series of scores. The approach proposed in this article belongs to this group. CFA of individual series of scores was proposed by von Eye (2002) as *unidimensional* CFA. In this approach, frequencies for a series of scores are estimated either to be constant, as in zero-order CFA, or with weights that account for a priori expected variations in a series. The difference to the approach proposed here is that, in unidimensional CFA, just as in all other models of longitudinal CFA, the expected cell frequencies are estimated based on a global representation of series of scores. Here, local representation is used.

Specifically, loess smoothing is used to estimate the expected frequencies for each observation. As was explained in the previous section on loess smoothing, the frequencies are estimated based on the same model for each time window. This model uses specifications made before analysis, that is, width of window, smoothing parameter, and degree of polynomial. However, the model remains the same over all time windows.

This approach is novel in the following two respects. First, instead of using the entire pool of data points, the estimation of expected frequencies uses only a selection of data points. Therefore, CFA types and antitypes indicate not only local deviations from expectancy but they are also local in the sense that the estimation of expected cell frequencies used

only local information. The range of information that is used depends on the width of the time window that is used for loess modeling. Second, types and antitypes do not indicate local contradictions to a global model. Instead, they indicate that a smoothing model that is a priori specified without using information from the data to be analyzed fails to describe the series within a particular time frame.

We call the approach proposed here *loess CFA*. It has the following characteristics:

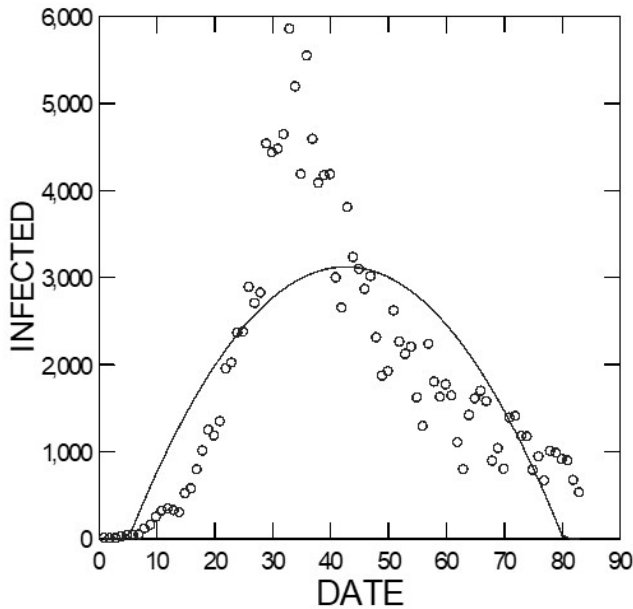
- it is nonparametric: no assumptions are made concerning the distribution of the data when the loess smoother is applied;
- it is a local regression approach: standard regression methods with least squares estimation are used for each time window (there exists a number of alternative regression models that can be considered as well);
- it, therefore, shares the characteristics of least squares regression such as sensitivity to outliers (note that robust loess smoothing approaches have been proposed);
- it is easily implemented, in particular when the observation points are equidistant (approaches to loess smoothing for non-equidistant observation points do exist, but they are computationally more intensive);
- it is more robust than most other time series smoothers;
- it requires relatively long series of observations (although there are no solid recommendations in the literature concerning the number of observations needed);
- the same significance tests for detecting types and antitypes can be used as in standard CFA (see von Eye & Wiedermann, 2021), and
- the same methods of α protection can be used as well; the number of tests is given by the number of observation points.

Data example. In the following paragraphs, we resume the analysis of the series of positive Covid-19 diagnoses that had been used for Figure 1. We proceed as follows. First, we illustrate that a global representation of these data is not satisfactory. Second, we illustrate the smoothing effects of time windows that differ in width. Third, we select a loess smoothing approach and subject the resulting expected frequencies to loess CFA.

To illustrate that global representation can be less than satisfactory, we just use a quadratic function. Figure 3 shows the result.

Figure 3 suggests that the global representation of the series of score by a quadratic curve describes the data poorly. The up and down is represented in part, the oscillations are not represented at all. This visual impression is supported by the non-significant correlation between observed and expected frequencies, $r = -0.147$ ($p = 0.184$). Adding to this model a linear trend increases the correlation to $r = 0.777$ ($p < 0.01$), but the data are still not well represented in the sense that there are large differences between expected and observed frequencies.

Figure 3
 Number of positive Covid-19 diagnoses in France in the 83 days after the first positive diagnosis; global quadratic smoother



We, therefore, move to loess smoothing. Here, we first illustrate that, whereas a wider window frame results in a smooth approximation curve, a narrower frame allows one to better capture the oscillations, but at the expense of a less smooth curve. Figure 4 displays the loess-smoothed curve under the following specifications:

- each time window contains a portion of 0.4 of all data points, thus allowing only long waves to be represented
- data points to the right and to the left are weighted inversely proportional to their distance from the target point by way of the tri-cube weight function
- quadratic local regression is used in each window.

Figure 4
 Number of positive Covid-19 diagnoses in France in the 83 days after the first positive diagnosis; loess smoother

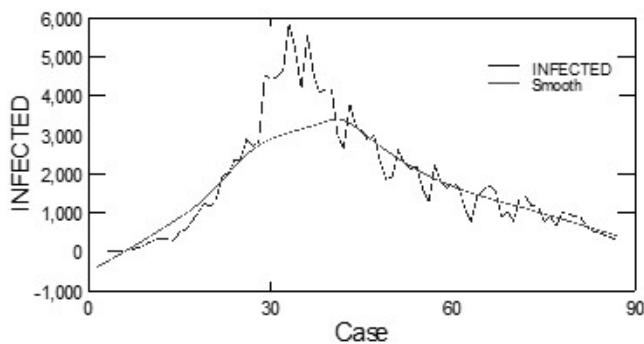
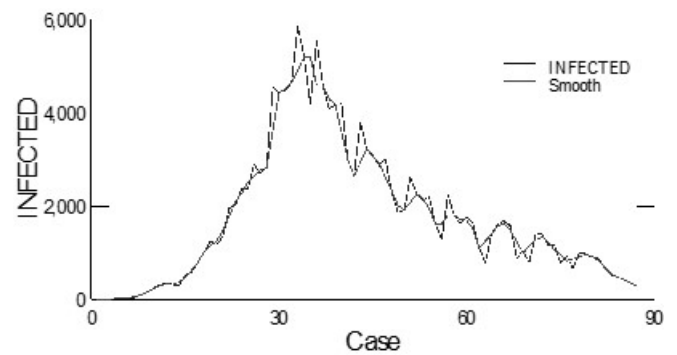


Figure 4 shows that the approximation of the observed frequencies with the loess smoother is clearly better than the approximation that is based on the global regression function.

The correlation between the observed and the expected frequencies now is $r = 0.92$, that is, close to perfect. Nevertheless, this result is not acceptable either, for two reasons. First, it estimates negative frequencies for the first five days. In the analysis of frequency data, negative estimates must be avoided. Second, it smooths over the oscillations. It was one of the goals of the presents analyses to capture at least some of the variability in the oscillations.

For these two reasons, we now select a narrower time window. Specifically, we reduce the portion of data points in the window to 0.05, thus allowing short-length waves of 2 Hz to be represented. All other model specifications are the same. Figure 5 displays the smoothed curve.

Figure 5
 Number of positive Covid-19 diagnoses in France in the 83 days after the first positive diagnosis; loess smoother with fine-grained resolution



Evidently, the loess-smoothed estimates now are much closer to the observed frequencies than in Figure 4. The correlation between the estimated and the observed frequencies now is $r = 0.983$, very close to perfect. Still, just 12 observed scores are exactly reproduced. This is desired because the smoothed curve is not more parsimonious than the observed curve when it reproduces all data points exactly.

Now, in spite of the extremely large correlation, Figure 5 shows that there are differences between the observed and the estimated progression of positive diagnoses. This is where CFA comes into play. CFA allows one to answer the questions of where in the series deviations are significant and how to interpret them. We take the four steps of CFA (von Eye & Wiedermann, 2021).

Step 1: Specification of base model. The key element of the interpretation of CFA types and antitypes is the CFA base model. In all applications of CFA thus far, the base model was defined as a probability model that specifies variable relations (cf. von Eye, 2004; von Eye & Gutiérrez Peña, 2004). In the present article, a new approach is proposed. Specifically, the base model is no longer expressed in terms of variable relations. Instead, the model is defined by the characteristics of the data generation process that is hypothesized to smooth the series. This process has, in the present case, three main characteristics.

First, it is unchanged over the entire time series. Second, simple quadratic polynomials are sufficient to describe the local shape of progression. Third, the time window used for the present analysis is relatively narrow. This last characteristic was specified so the oscillating characteristics of the series could be captured. This was deemed necessary because there are systematic changes over the days of the week. These changes reflect characteristics of data collection, that is, on Sundays, fewer tests and diagnoses are performed than on other days of the week. This results in day of the week-related oscillations. Figure 4 suggests that the selected model is capable of approximating these oscillations, up to a point.

Unchanged from standard CFA, the interpretation of CFA types and antitypes is conducted with reference to the CFA base model. Here, types and antitypes do not suggest that variable relations exist that were not included in the base model. Instead, types and antitypes suggest that the model that was used to represent the local progression is, for a particular target point and its neighborhood, insufficient. In most cases, one can assume that the model is not complex enough. Here, higher order polynomials can be considered. Substantively, one can ask what events caused the locally

increased complexity of the progression.

Step 2: Significance testing. The numbers of positive Covid-19 diagnoses are large. Therefore, we can use the z-test and protect α with the Bonferroni procedure. The protected significance threshold becomes $\alpha^* = 0.05/83 = 0.0006$. This value corresponds to an absolute z-score of 3.238. Observed z-scores more extreme than this value point at target points that constitute types or antitypes.

Step 3: Performing CFA. Table 1 displays the results of CFA. The first column in the table contains the days of observations, beginning with the first day of a positive diagnosis. The second column contains the observed counts. The third column contains the frequencies that were estimated under the wide window (Figure 4; Loess04). The fourth column contains the frequencies that were estimated under the narrower window (Figure 5; Loess005).

The values in these two columns correlate to $r = 0.929$. Again, this is an extremely high correlation. Still, the narrower window results in a much better approximation. In particular, Table 1 shows that all estimates under the narrower window (Loess005) are positive - a must for frequency data.

Table 1

Loess-CFA of the series of Covid-19 diagnoses in the first 83 days of the pandemic in France, 2020

Date	Weekday	Infected	Loess04	Loess005	z-value	Type/ Antitype?
1	Friday	1	-393.46	1.12	0.11	-
2	Saturday	1	-313.09	0.71	0.34	-
3	Sunday	0	-231.16	5.73	2.39	-
4	Monday	19	-147.6	17	0.49	-
5	Tuesday	31	-62.41	29.28	0.32	-
6	Wednesday	37	24.39	35.86	0.19	-
7	Thursday	39	112.77	59.09	2.61	-
8	Friday	111	202.65	101.85	0.91	-
9	Saturday	151	293.97	166.48	1.2	-
10	Sunday	245	386.64	238.14	0.44	-
11	Monday	315	480.59	302.69	0.71	-
12	Tuesday	342	575.71	327.67	0.79	-
13	Wednesday	319	671.89	318.72	0.02	-
14	Thursday	295	769.03	366.95	3.76	A
15	Friday	516	867.2	468.91	2.17	-
16	Saturday	571	967.48	617.69	1.88	-
17	Sunday	788	1073.89	787.72	0.01	-
18	Monday	1004	1209.12	1010.68	0.21	-
19	Tuesday	1244	1352.32	1152.34	2.7	-
20	Wednesday	1176	1504.35	1238.98	1.79	-
21	Thursday	1341	1666.09	1466.73	3.28	A
22	Friday	1947	1837.83	1776.53	4.04	T
23	Saturday	2016	2018.5	2087.31	1.56	-
24	Sunday	2359	2202.78	2267.06	1.93	-
25	Monday	2370	2381.22	2496.21	2.53	-
26	Tuesday	2885	2544.28	2625.3	5.07	T
27	Wednesday	2699	2686.19	2762.35	1.21	-
28	Thursday	2819	2808.17	2819	0	-
29	Friday	4532	2913.4	3624.5	15.07	T
30	Saturday	4430	3003.02	4430	0	-
31	Sunday	4472	3078.38	4507.94	0.54	-
32	Monday	4639	3141.84	4639	0	-
33	Tuesday	5851	3196.68	4913.5	13.37	T
34	Wednesday	5188	3245.91	5188	0	-

35	Thursday	4181	3292.63	5188	13.98	A
36	Friday	5542	3339.9	4585	14.13	T
37	Saturday	4585	3391.34	4585	0	-
38	Sunday	4079	3448.52	4289.32	3.21	-
39	Monday	4168	3504.9	4168	0	-
40	Tuesday	4180	3539.75	3580.5	10.02	T
41	Wednesday	2993	3525.2	2993	0	-
42	Thursday	2646	3451.39	2646	0	-
43	Friday	3800	3335.8	2936.5	15.93	T
44	Saturday	3227	3205.26	3227	0	-
45	Sunday	3089	3075.59	3063.67	0.46	-
46	Monday	2860	2950.04	2905.61	0.85	-
47	Tuesday	3008	2829.08	2619.52	7.59	T
48	Wednesday	2305	2712.27	2334.02	0.6	-
49	Thursday	1865	2598.61	2010	3.23	-
50	Friday	1916	2487.98	1916	0	-
51	Saturday	2615	2381.28	2086	11.58	T
52	Sunday	2256	2279.94	2256	0	-
53	Monday	2114	2185.09	2165.35	1.1	-
54	Tuesday	2197	2096.6	1947.55	5.65	T
55	Wednesday	1616	2013.41	1616	0	-
56	Thursday	1287	1934.13	1616	8.18	A
57	Friday	2230	1858.59	1796	10.24	T
58	Saturday	1796	1788.07	1796	0	-
59	Sunday	1622	1723.5	1715.87	2.27	-
60	Monday	1765	1664.43	1690.56	1.81	-
61	Tuesday	1637	1609.65	1511.59	3.23	-
62	Wednesday	1100	1557.72	1105.4	0.16	-
63	Thursday	789	1507.13	1248.95	13.01	A
64	Friday	1412	1457.17	1400.21	0.32	-
65	Saturday	1603	1408.42	1574.53	0.72	-
66	Sunday	1690	1361.99	1635.8	1.34	-
67	Monday	1572	1318.4	1464.13	2.82	-
68	Tuesday	888	1279.71	1239.05	9.97	A
69	Wednesday	1032	1240.64	990.26	1.33	-
70	Thursday	795	1201.09	1143.1	10.3	A
71	Friday	1383	1161.32	1273.16	3.08	-
72	Saturday	1403	1121.53	1336.85	1.81	-
73	Sunday	1173	1081.79	1235.82	1.79	-
74	Monday	1169	1042.13	1063.13	3.25	T
75	Tuesday	782	1002.52	951.79	5.5	A
76	Wednesday	935	962.95	829.42	3.67	T
77	Thursday	659	923.42	867.4	7.08	A
78	Friday	999	883.92	905.55	3.11	-
79	Saturday	980	844.49	965.07	0.48	-
80	Sunday	907	805.13	923.42	0.54	-
81	Monday	892	765.82	830.59	2.13	-
82	Tuesday	667	726.55	690.49	0.89	-
83	Wednesday	525	687.27	555.62	1.3	-

Step 4: Interpretation of types and antitypes. Table 1 shows that nine antitypes and 13 types emerged. Of these, we first interpret the most extreme ones. The most extreme type is constituted by Day 43. 3800 positive diagnoses were counted, but the loess smoother made one expect only 2936.5 cases. The hypothesized data generation mechanism was clearly unable to capture this strong upswing. The most extreme antitype is constituted by Day 35. 4181 positive diagnoses were counted for this day, but 5188 had been expected. Here, the hypothesized data generating mechanism was unable to capture the downswing in the progression. This applies accordingly to the other types and antitypes.

Second, we interpret the resulting pattern of types and antitypes. Interestingly, seven of the nine antitypes are

constituted by days that are multiples of seven. In addition, these seven antitypes all occur on Thursdays. This is significantly more than expected under the assumption of a uniform distribution of antitypes over the days of a week ($p = 0.017$). Considering that Santé Publique did not publish any discussion of diagnose patterns or trends, we are in search of an explanation of this result. Among the possible interpretations is the following one. At the beginning of the pandemic, it took about four days before the results of the Covid-19 tests were available. The Thursday antitypes may, therefore, reflect the fact that, on the Sundays before, fewer tests were administered than on other days. The CFA base model of loess smoothing seems to be sensitive to these drops.

Types tend to occur more often than average on Fridays.

Five of the 13 types occurred on a Friday. This, however, is a non-significant deviation from expectancy. Longer series of data may be needed to determine whether there is a connection of these increases with the drops on Thursdays.

In all, we conclude that the rhythm that was hypothesized for the days of the week was much better captured under the narrower than under the wider window, but still not well enough to prevent these seven antitypes and 12 types from emerging. There seems to be a significant pattern only for the antitypes. In this respect, the downswings resulted in a pattern that is harder to capture by the hypothesized data generation mechanism than the pattern of the upswings. Overall, however, more upswings than downswings were missed. It needs to be determined whether this reflects the progression of the pandemic, day of the week-related characteristics of testing, or both.

Discussion

In this article, we propose a new approach to CFA. The new approach differs from existing approaches to CFA in two major aspects. First, thus far in CFA, the logic of searching for types and antitypes consisted of specifying a base model that included all effects, e.g., variable relations, that are not of interest to the researchers. Types and antitypes then suggest where in the data space the effects exist that are of interest to the researchers. The new approach follows this logic. However, instead of specifying a base model of variable relations, it specifies a base model in terms of a data generating process. Types and antitypes then suggest the sector in the data space in which a different or more complex data generation process must be at work.

The second major difference between CFA as it was known so far and the new approach is that the new approach is tailored to be applied to series of individual observations, e.g., time series (standard CFA approaches for longitudinal data rely on aggregated counts). To approximate a series, a data generating process is hypothesized that explains the progression locally, that is, within the intervals of an a priori specified width.

Just as in earlier approaches to CFA, the number of options to explain the emergence of types and antitypes is limited. In CFA as it was known so far, these options consist of variable relations. For example, in first order CFA, these options include all possible variable interactions. In Prediction CFA (P-CFA), these options include all possible predictor-criterion relations. Here, the options concern the three main elements of loess smoothing, the width of the window, the function used to locally approximate progression, and the weights used for the neighbors of the target point. In the example in the last section, we moved from Figure 4 to Figure 5 by leaving the function and the weights untouched. However, we narrowed the time window. This way, we opened the door to a finer-grained resolution of the smoothing process, and we know that the differences between Figures 4 and 5 are solely due to the narrowed window.

In the analysis of time series, a large number of methods

has been proposed. A thorough comparison of these methods is beyond the scope of this article. However, it can be said that, some of these, e.g., spline smoothing, can result in very similar results as loess smoothing. In addition, spline smoothing requires less calculation effort than loess smoothing. For the current purposes, loess smoothing seems more useful nevertheless, for two reasons. First, loess smoothing can be tailored so that both the fast and the slow changes are modeled separately. Second, it is more local in the sense that changes within individual and narrow windows can be reflected. This characteristic is of particular importance for CFA. Still, the configural analysis of time series that is based on curves that describe the data generation process in just one function can be an interesting option, one that will be explicated in future work.

Generalizations of the proposed new CFA approach are straightforward. The most important could be the generalization to loess-smoothing of more than one variable. This way, in time series, progression hyperplanes would be smoothed instead of progression lines. This is promising but it is material for future work.

Other generalizations concern the observation points. In the description of the method and in the example given in this article, the observation points were equidistant. In many empirical studies, equidistant observations are hard to realize. For example, in school settings, series of observations are interrupted because children are on vacation. In psychotherapy, time intervals between therapy sessions can vary depending on the needs of patients, or WiFi reception varies because of interferences. Methods for loess-smoothing for non-equidistant observation points exist and can be incorporated in the loess CFA framework.

A third way of generalization of the method proposed here consists of incorporating the new method in the context of existing CFA methods. It is, for example, conceivable, to incorporate loess CFA in the context of moderator CFA. Moderator variables would then be used to explain variations of type/antitype patterns across individuals or groups of individuals. In a similar way, smoothed curves could be generated and subjected to CFA for periods before and after certain interventions. This, again, is material for future work.

Author contributions

The three authors contributed equally to this article, and all consent to the publication of this work.

Declaration of interests

The authors declare that there are no conflicts of interests.

Action editor

Lars-Gunnar Lundh served as action editor for this article.

Open access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License

(<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Cleveland, W.S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74, 368, 829-836.
<https://doi.org/10.1080/01621459.1979.10481038>
- Cleveland, W.S., & Devlin, S.J. (1988). Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, 83, 403, 596-610. <https://doi.org/10.1080/01621459.1988.10478639>
- Dagum, E.B., & Luati, A. (2000). Predictive Performances of Some Linear and Nonlinear Smoothers for Noisy Data. *Statistica, LX*, 635 - 654. <https://doi.org/10.6092/issn.1973-2201/1157>
- Dagum, E.B., & Luati, A. (2001). A Study of the Asymmetric and Symmetric Weights of Kernel Smoothers and their Spectral Properties. *Estadística, Journal of the Interamerican Statistical Society, Special Issue on Time Series Analysis*, 53, 160-161.
- Dagum, E.B., & Luati, A. (2003). Global and local statistical properties of fixed-length nonparametric smoothers. *Statistical Methods & Applications*, 11, 313-333.
- Lienert, G. A. (1968). *Die "Konfigurationsfrequenzanalyse" als Klassifikationsmethode in der klinischen Psychologie*. Paper presented at the 26. Kongress der Deutschen Gesellschaft für Psychologie in Tübingen 1968.
- Santé Publique France (2020) COVID_19: point épidémiologique du 10 mars 2020. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-10-mars-2020>. Retrieved on 3/10/2020.
- Savitzky, A., & Golay, M.J.E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 8, 1627-1639. <https://doi.org/10.1021/ac60214a047>
- von Eye, A. (2002). *Configural Frequency Analysis - Methods, Models, and Applications*. Mahwah, NJ: Lawrence Erlbaum.
- von Eye, A. (2004). Base models for Configural Frequency Analysis. *Psychology Science*, 46, 150-170.
- von Eye, A., & Gutiérrez Peña, E. (2004). Configural Frequency Analysis - the search for extreme cells. *Journal of Applied Statistics*, 31, 981-997.
<https://doi.org/10.1080/0266476042000270545>
- von Eye, A., Mair, P., & Mun, E.-Y. (2010). *Advances in Configural Frequency Analysis*. New York: Guilford Press.
- von Eye, A., & Wiedermann, W. (2021). *CFA. Configural Frequency Analysis*. New York: Springer (in preparation).
- Wiedermann, W., Herman, K. C., Reinke, W., & von Eye, A. (2021). Configural frequency trees. Development and Psychopathology. (in press)
- Wilkinson, L., Blank, G., & Gruber, C. (1996). Desktop data analysis with SYSTAT. Upper Saddle River, NJ: Prentice Hall.