Chapter 14
Investigating the Long Memory Process in Daily High School Attendance Data

Matthijs Koopmans

Complex dynamical systems research is motivated by a desire to understand how systems maintain stability over the longer term, and how they transform themselves. To that end, the early cybernetic literature has maintained that the role of time needs to be considered when trying to establish a causal connection between outcomes and input conditions (Ashby, 1957; Wiener, 1961). While the causal attribution of outcomes to changing input conditions is part and parcel of many educational studies, there have been few attempts to deliberately model time when establishing this causality (Koopmans, 2014a). The description of large samples of sequentially organized data through time series analysis is quite common in many other disciplines, such as cardiology (heart rates), meteorology (temperature, precipitation), and econometrics (mortgage rates, interest rates), and in fact, time series can be found on an almost daily basis in newspapers such as the *New York Times* and the *Wall Street Journal*.

Time series are useful whenever it needs to be estimated whether the passage of time influences the causal mechanisms that predispose systems to behave in a certain way. They have been used to study phenomena as diverse as irregular heartbeat (Peng et al., 1993), blood cell perfusion in rat brains (Eke et al., 2000), seasonal variability in the teen pregnancy rates in the state of Texas (Hamilton, Pollock, Mitchell, Vincenzi, & West, 1997), and much more. In spite of the fact that the conceptual foundations of this approach for education have been lucidly laid out quite a long time ago (Glass, 1972), the use of time series in education has not received as much attention as one might expect given the time dependency of many of the processes of interest to the discipline: students learn over time, achievement gaps get narrowed over time, and teachers manage time when they plan and execute their lessons. This lack of attention to the time aspect reflects a tendency, particularly

M. Koopmans (✉)
Mercy College, School of Education, Mercy Hall, Rm. SE 31, 555 Broadway,
Dobbs Ferry, NY 10522, USA
e-mail: mkoopmans@mercy.edu

M. Koopmans, D. Stamovlasis (eds.), *Complex Dynamical Systems in Education*, DOI 10.1007/978-3-319-27577-2_14

in applied research circles, to build studies around the prediction of educational outcomes at the group level, rather than the underlying dynamics of educational processes at the individual level (Koopmans, 2014a).

Cross-sectional methodologies rely on the tacit assumption that measurement results obtained over a large sample of cases can be generalized across a large time spectrum, as defined by the scope of the conclusions drawn from those measurements. Does a "snapshot" standardized test result characterize stable achievement levels over the scope of, say, an entire school year? The assumption that you can generalize from cross-sectionally obtained group averages to the entire time spectrum for individual group members is known in the literature as the *ergodic assumption* (Molenaar, 2004; Chap. 8). If this assumption cannot be taken for granted, its verification becomes an empirical issue, requiring a detailed analysis of individual cases. While Molenaar originally made this argument in the context of psychological research, similar argument can be made for education, which, in many ways relies on the same measurement practices for statistical inference, i.e., the measurement of behavioral constructs across groups of individuals (Kerlinger, 1977). An interesting question to contemplate for educational researchers is what we can learn about systemic behavior and transformation thereof from the detailed and statistically rigorous analysis of the contribution of time to behavior in individual cases. Such understanding cannot be easily obtained through conventional linear statistical techniques, which typically rely on the aggregation of the findings across individuals for statistical inference (Neter, Wasserman, & Kutner, 1985).

The underlying assumption when using time-invariant measures is that the systems under study are stable. This assumption of stability has also pervaded the early dynamical literature that traditionally assumed that systems were in principle in a state of equilibrium, except for the instability that accompanies a transformation process (e.g., Lewin, 1947). The more recent literature on dynamical processes has challenged this assumption and proposed that healthy systems may often be in a state of disequilibrium (Bak, 1996; Goldstein, 1988), resulting in an openness to transformation (Stadnitski, 2012b), whereas this proclivity may be absent in systems that are stable in the sense that their behavior is highly predictable based on past occurrences.

Moreover, systems may appear stable for long periods of time while the endogenous process brings those systems to a critical state called *self-organized criticality* (Bak, 1996). The prototypical example of self-organized criticality is the sand pile model, which states that a continued supply of sand to a pile on a flat surface causes occasional avalanches that reorganize the pile, ostensibly to reduce the friction between the grains that result from the accumulation (Jensen, 1998). The state of friction in a system where change is imminent is called self-organized criticality or being "at the edge of chaos," and it is seen as an indicator of systemic complexity (Waldrop, 1992). One implication of the idea of self-organized criticality is that there is a continuous relationship between the small ordinary events that define the endogenous process in the system and the large cataclysmic events that produce transformation in the system, requiring a single analytical framework capturing both aspects.

One important characteristic of self-organized criticality in systems is self-similarity, also referred to as fractality, $1/f^{\beta}$noise or pink noise. A well-known example of self-similarity is the coastline of Norway, which on a small scale replicates patterns that are also observed on a large scale (Feder, 1988), although the scale at which they replicate is not constant. This independence of the patterns observed on the scale at which they are observed is called *scale invariance*. When measurements are conducted over time, patterns of variability can similarly replicate themselves. Such self-similarity occurs when the same variability patterns are observed within an undetermined variety of different time frames, suggesting an alternating but unpredictable pattern of stability/instability.

Complexity and nonlinear dynamical system theories provide a rich array of transformative scenarios, such as bifurcation and period doubling, sensitivity to initial conditions, hysteresis, second-order change, coupled oscillators, and change through self-organized criticality (Koopmans, 2009), and the search for empirical manifestations of those scenarios requires the detailed analysis of sequentially ordered observations in almost all instances. While time series is a common statistical technique, its fine-tuning to specifically address transformative hypotheses put forward in the dynamical literature is a relatively recent development. Two aspects that have generated particular analytical interest are the use of time series analysis to detect sensitivity to initial conditions and chaos (Kantz & Schreiber, 2004; Kaplan & Glass, 1995; Sprott, 2003), and the measurement of self-organized criticality, fractality, and long memory processes (Beran, 1994). The analysis presented here focuses on the latter of these two applications.

School Attendance as a Dynamical Process

Few educators would dispute that attending school is critical to successful educational outcomes, as it is a prerequisite to exposure to classroom instruction and the learning opportunities it provides. In addition, school attendance is also a mediating variable in the system of causal relationships that includes parental support, student academic engagement, instructional effectiveness, and academic attainment (Astone & McLanahan, 1991; Balfanz & Byrnes, 2012; Kemple, Segeritz, & Stephenson, 2013; Kemple & Snipes, 2000; Roby, 2003). In spite of its apparent importance, the analysis of school attendance has taken the backseat to outcomes such as academic achievement, high school dropout, and college persistence behavior, and to the extent that attendance data get reported, it is reported in aggregated form, averaging daily attendance rates over weekly, monthly, or yearly periods (see, e.g., National Center of Education Statistics, 2008), requiring us to assume that those rates are stable over time. Reporting attendance aggregated across the time spectrum results in significant information loss. A time-sensitive view of attendance may help reveal how existing attendance rates impact future attendance over the immediate and longer term, whether there are cyclical patterns to this impact, and what the timing might be of the response of attendance rates to external events or conditions.

An opportunity presented itself to conduct a statistically rigorous analysis of the dynamical processes that may be manifest in educational time series when the New York City Department of Education started recording and publishing the daily attendance rates of all of its schools in 2004, and continued to do so up to the day of this writing. The resulting data sets provide highly detailed information to estimate about how attendance behavior is affected by the progression of time, how attendance patterns differ from one school to the next, and to what extent transformative scenarios such as the ones mentioned above play out over these attendance trajectories.

Most teachers and school administrators are probably well aware of the ebbs and flows in the daily attendance in their classrooms and school buildings. In formal research, these fluctuations get obfuscated by the aggregations that are seen as necessary to summarize the data meaningfully. Hence, the findings of this research do not connect effectively to local knowledge in the schools about daily attendance (Koopmans, 2015). A related point is that applied research in education tends to prefer the cross-sectional estimation of complex cause-and-effect relationships instead of the estimation of the endogenous process through which those relationships are generated (Sulis, 2012). As a result, the literature provides little guidance about what to expect with regard to the short-range dependencies in daily attendance rates, nor the correlations between observations over longer time periods. The work discussed here aims to address this gap.

Using Time Series Analysis to Uncover Dynamical Patterns

The purpose of the analysis presented here is to uncover the dynamical patterns in daily attendance rates, and illustrate why the estimation of those patterns may yield relevant insights into attendance behavior at the school level. Data were obtained from a total of seven schools and some data preparation was done to make the information suitable for a time series analysis. Since such analyses do not permit missing values, a nearest-neighbor imputation was conducted in instances when daily attendance was not recorded on three or fewer subsequent occasions in a given week. If more than 3 days were missing from a given week, that week was removed in its entirety from the series. Similarly, the summer and winter recess was not considered and the last session before and first session after recess were connected as neighbors to ensure the integrity of the dynamics of the temporal ordering of the information.

The two sections that follow will first describe how the estimation of short-range error dependencies (autocorrelation) proceeds in a conventional autoregressive integrated moving average (ARIMA) analysis (Box & Jenkins, 1970; Cryer & Chan, 2008). It is then shown how long-range patterns can be estimated through an extension of this framework called autoregressive fractionally integrated moving average (ARFIMA), a method introduced by Granger and Joyeux (1980) and Hosking (1981) to model the slowly decaying autocorrelations that characterize the long-term memory process. A third section describes the use of power spectral analysis, a procedure used to convert time series plots into plots that show the periodicity of the data. This analytical procedure can be used to detect long-term fractal patterns (Delignières, Torre, & Lemoine, 2005; Wagenmakers, Farrell, & Ratcliff, 2004).

The capability of the combined ARIMA/ARFIMA approach to address both short- and long-range error dependencies within a single analytical framework (Wagenmakers et al., 2004) makes the approach particularly attractive to analyze daily attendance rates, and sets it apart from many other approaches to the detection of fractality, such as power spectral density (PSD) analysis (Eke et al., 2000), de-trended fluctuation (DFA) analysis (Peng et al., 1993), and rescaled range (R/S) analysis (Hurst, 1965), none of which is particularly well suited to differentiate short-range and long-range processes. Delignières et al. (2005) provide a lucid overview of these and related approaches.

Like ARFIMA, Thornton and Gilden's (2005) spectral likelihood classification is designed to distinguish short-term from the long-term processes, but it approaches the issue as an "either/or" proposition; that is, the short-term model and the long-term model compete to provide the best fit to the data. As a result, this approach does not enable the investigator to examine the contribution of long-range processes to the variability in the trajectory *over and above* the contribution of the short-range ones. Thornton and Gilden rightly argue that such as assessment is unlikely to be of great theoretical interest when first-order dependencies (i.e., correlations between neighboring values on the trajectory) are at issue, but in the context of the analysis of daily attendance patterns in high schools, the question is pertinent whether the long-range modeling component needs to be supplemented by seasonal estimators, i.e., short-range features that are of substantive importance to the field such as the days of the school week. Our knowledge about seasonal fluctuations in attendance may facilitate planning at the classroom, school building, and policy level, and may help us better understand the interplay between exogenous (e.g., parents, SES) and endogenous influences (i.e., school attendance rates in the near and distant past). Such estimation may also enhance our understanding about the extent to which the prediction of variability in daily attendance trajectories is relatively straightforward and to what extent it requires dealing with the complexities in the system's behavior. The ARFIMA approach is better equipped to make these distinctions than approaches based on power spectra. However, the particular strength of power spectral analyses compared to ARFIMA is that the former procedure does not require any assumptions with regard to the distribution of observations across the spectrum. Specifically, it can reliably estimate fractality regardless of whether the original time series is stationary or not, whereas ARFIMA requires stationary data (Stadnitski, 2012b; Wagenmakers et al., 2004), i.e., data whose statistical properties are constant across the entire time spectrum.

Short-Range Dependencies

In this section, I'd like to discuss the estimation of short-range dependencies, a statistical procedure that has been part and parcel of conducting time series analysis for many decades now. Let us start with an example. Figure 14.1a shows the daily

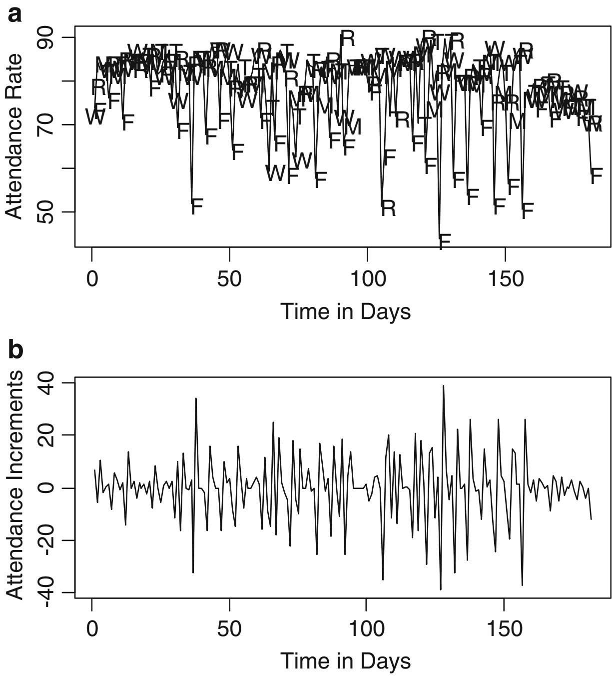

Fig. 14.1 (**a**) Daily attendance 2009–2010 in School 1 with the days of the week marked ("R" represents Thursday, $N = 183$); (**b**) first difference $Y_t - Y_{t-1}$ of the attendance rates in School 1 (attendance increments)

attendance for the entire academic year 2009–2010 in one New York City high school (School 1), marking the days of the week ("R" represents Thursday). The trajectory displays a somewhat drooping appearance with many outlying observations falling way below what would be the average of the series. It can also be seen that there is an overrepresentation of Fs (Fridays) among those low-lying observations. The implication of this pattern would be that average daily attendance rates aggregated across the time spectrum systematically overestimate attendance on Friday and underestimate attendance on the other days of the week.

Figure 14.1b shows a trajectory of attendance increments, or first differences ($Y_t - Y_{t-1}$) in that same school. It can be seen in that figure that in the course of the school year, the differences between given observations and their immediately preceding neighbors become larger, resulting in increased variability, which is to say that the trajectory shows heteroscedasticity across the time spectrum. This trend would go unheeded if traditional central tendency and variability measures are used to characterize these data, leaving us unaware of the increased turbulence in daily attendance as the year progresses.

These examples illustrate very clearly why measures of central tendency and variability are insufficient to characterize the distribution of daily attendance data, as these measures ignore the skewness and the cycles in the first example, and they ignore the increased variability over time in the second one. These examples also indicate why conducting an ordinary least squares regression of daily attendance rates on time yields a biased estimate of their relationship. The observations are not independent, as shown in the first example, and the assumption of homoscedasticity is violated in the second example. The failure of these traditional estimates to handle characteristics that are typical of time-dependent data is part of what motivates ARIMA, which is designed to distinguish two types of error dependency: *the autoregressive process* (AR), and the *moving average process* (MA).

The AR model predicts the value of Y_t as a linear combination of its own past values, plus an error term that is presumed to be an independent identically distributed random variable. The MA model predicts Y_t in terms of accumulated error disturbances, also called innovations. Appendix 1 explicates the ARMA models formally. The investigator can control the number of lags that are used in this prediction for each of these two modeling components. To ensure an unbiased estimation of AR and MA processes, it is essential to verify the stationarity assumption, i.e., the constancy of statistical properties of the data across the entire trajectory. In case of non-stationarity, the first difference of the time series $(Y_t - Y_{t-1})$ is typically used for the estimation. A process that requires such differencing to estimate the ARMA components is called an integrated ARMA or ARIMA process (Cryer & Chan, 2008).

There is a variety of ways to test for the stationarity of a time series. The most well known is the augmented Dickey-Fuller (ADF) test (Fuller, 1996), which regresses the first difference of an observed time series on lag 1 or the original series, and on the past k lags of the first difference of the series. It is then tested whether the beta coefficient in the regression model associated with the lag 1 observation is different from zero, using the parameters for the past k lags as covariates. Rejection of the null hypothesis confirms stationarity of the series (Cryer & Chan, 2008). Thus, using conventional notation, ARIMA (p, d, q) defines the number of AR parameters p and the number of MA parameters q included in the estimation process. The parameter d refers to the order of differencing required, i.e., $d = 0$ for the stationary process, and $d = 1$ for the use of the first difference of a non-stationary process.

Figure 14.2a, b shows simulated examples of a stationary and a non-stationary time series for a sample of 180 observations. Figure 14.2c shows the first difference of the trajectory in Fig. 14.2b, which results in stationarity. In the series shown in Fig. 14.2a, c, it can be seen that the patterns of variability look pretty similar across the series and that the mean of zero appropriately characterizes its central tendency. This is clearly not the case for the trajectory shown in Fig. 14.2b, which characterizes non-stationarity. This latter simulation shows what is known as a *random walk* or *Brownian motion*, an unstable system with strongly correlated observations. The results of the ADF test on these three trajectories are as follows: $\text{ADF} = -4.95$, $p < 0.01$; $\text{ADF} = -1.43$, $p > 0.01$; and $\text{ADF} = -5.95$, $p < 0.01$ for the series in Fig. 14.2a, b, and c, respectively, using $k = 4$ as the lag order. These results confirm the properties that these simulations were set out to show.

When estimating short-range effects across the time spectrum, it is often productive to inspect to the autocorrelation function plots to detect the patterns of

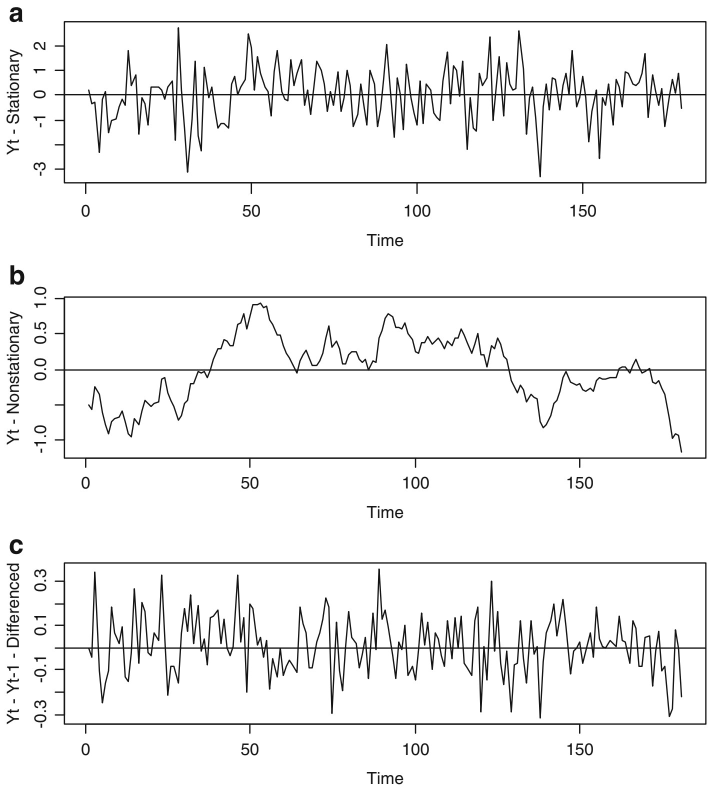

Fig. 14.2 (**a**) Simulation of a stationary time series ($N = 180$, ADF $= -4.95$, $p < 0.01$); (**b**) a non-stationary time series ($N = 180$, ADF $= -1.43$, $p > 0.01$); and (**c**) the first difference $Y_t - Y_{t-1}$ of the series in (**b**) produces stationarity (ADF $= -5.95$, $p < 0.01$)

dependency residing in the data. The use of these plots is illustrated in Fig. 14.3. Three simulated trajectories ($N = 180$) are shown in the left panels and the corresponding ACF plots are shown on the right. Figure 14.3a shows a simulated trajectory without error dependencies (white noise). In this situation, knowing the trajectory does not improve our ability to predict subsequent observations. The ACF plot corresponding to this situation is shown on the right. The spikes in the plot indicate the size of the autocorrelations at the lags indicated on the abscissa. The dotted lines indicate the 95 % confidence interval. The plot shows that none of the autocorrelations up to lag $k = 30$ are different from zero. The trajectory in Fig. 14.3b shows the clustering of neighboring observations that comes with autocorrelation, giving the trajectory in its entirety less of a random appearance than the one shown in Fig. 14.3a. An autoregressive process was generated using an AR (1) model with $\varphi = 0.70$, also with 180 observations. The ACF plot shows what a positive AR (1) process typically looks like. The correlations at the first few lags are significantly different from zero, but they rapidly recede to non-significance as the lag order

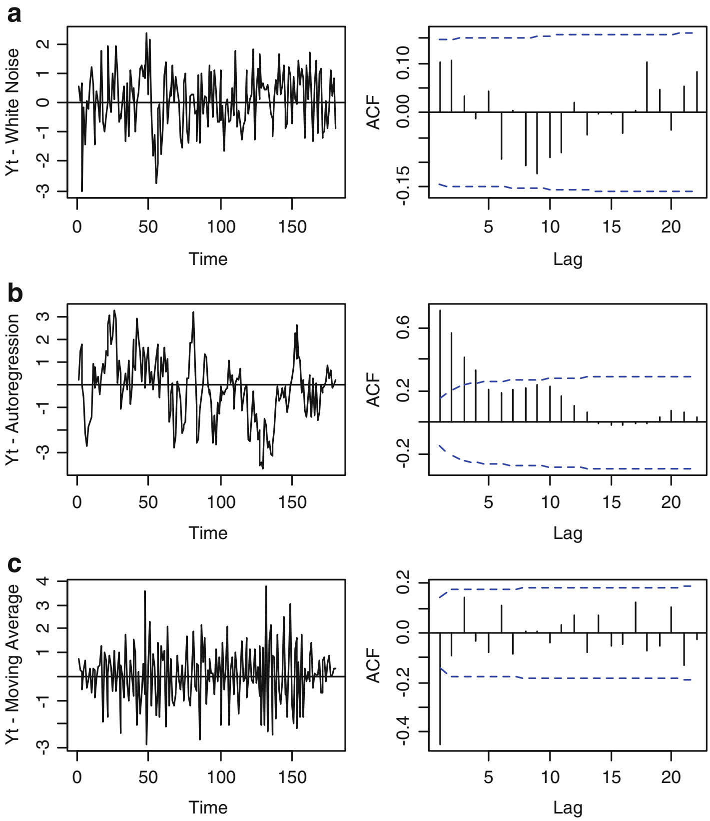

Fig. 14.3 Three simulated time series (*left* panels) with corresponding ACF plots (*right* panels). (**a**) White noise; (**b**) autoregression ($\varphi = 0.70$); (**c**) moving average ($\theta = 0.70$). $N = 180$

increases. Figure 14.3c illustrates an MA (1) scenario at $\theta = 0.70$. A different clustering pattern can be observed in this latter series where consecutive observations tend to alternate across the mean of zero, as is also indicated by the negative autocorrelation shown in ACF plot for the first lag. Note also that, typical of the moving average process, after the first spike, the autocorrelations immediately recede to non-significance at subsequent lag values. The examples presented here can be extended to include AR and MA processes at negative parameter values, multiple AR (p) or MA (q) parameter values, and combined ARMA (p, q) estimates (see, e.g., Box & Jenkins, 1970; Cryer & Chan, 2008; Shumway & Stoffer, 2011).

Seasonal ARMA Processes

One of the advantages of the ARIMA/ARFIMA approach is that the number and size of the lags included in the predictive models are fully up to the investigator, and there may be substantive reasons to model predictions around particular lag sizes, such as cycles denoting the days of the week or months in a year. For the analysis of school attendance in particular, the 5 days of the week are of particular interest to estimate whether daily attendance rates have a seasonal cycle. Consequently, over and above the estimation of the impact of immediately neighboring values (i.e., attendance on the previous day or 2 days), as illustrated above, we would like to estimate the impact of last week's attendance rate. Does knowing the attendance rate on a given day of the week improve our prediction of attendance on that same day the following week? The trajectory shown in Fig. 14.1a illustrates the relevance of this estimation. Appendix 1 shows the formal modeling features of the seasonal ARMA model.

An empirical example of the weekly cycles is shown in Fig. 14.4a, which shows the daily attendance trajectory for School 2 in the 2009–2010 school year, as well as the ACF plot. While the cyclical dependencies may be difficult to detect in the time series, the ACF plot brings them out very clearly as a pronounced spike at the fifth lag. This ACF plot also points to the absence of short-range dependencies at other lag values. Figure 14.4b shows the residuals of the ARIMA model that successfully models the seasonal dependency at five lags ($\varphi_1 = 0.90$, $\theta_1 = -0.71$, and $\theta_2 = 0.16$). The trajectory on the left suggests randomness, and the ACF plot confirms that there are no remaining short-range dependencies in the data. The extreme values shown in Fig. 14.4a were modeled using an intervention analysis framework (Cryer & Chan, 2008, see Koopmans, 2011 for further details about that aspect of the analysis).

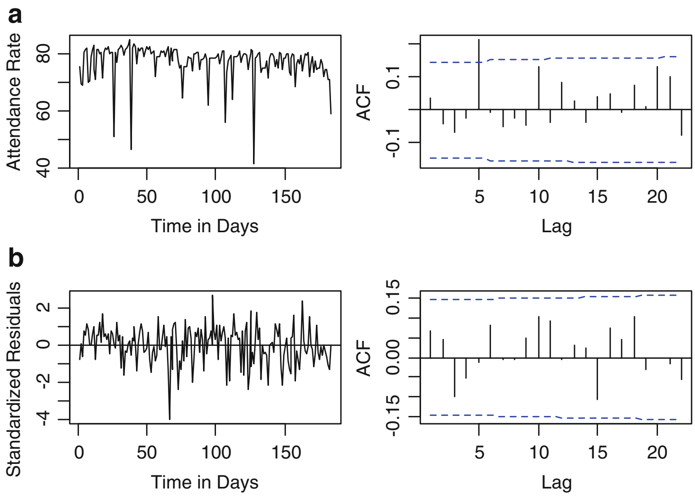

Fig. 14.4 (**a**) Trajectory of daily attendance rates in School 2 (*left* panel) with the corresponding ACF plot at the *right*; (**b**) residuals of a successful ARIMA model with corresponding ACF plot. $N = 183$

Long-Range Dependencies

The estimation of long-range dependencies helps determine whether there is evidence of self-organized criticality in the trajectories. Self-organized criticality would indicate that, as in the sand pile experiments discussed above, there are instances of critical instability and a repeating tension-release process in the face of continued input. In the data discussed here, perhaps long episodes of required attendance behavior create the need for incidental release, with a state of self-organized criticality immediately preceding this release. In time-sensitive measurements, indicators of self-organized criticality are the presence of self-similar patterns, and strong autocorrelations over a wide time spectrum. An important part of the data for long-range dependencies therefore is the detection of these two patterns.

Self-similarity refers to the replication of certain patterns at various scales, i.e., patterns within patterns. These patterns do not replicate in a strictly deterministic way. Rather, it is their *general impression* that remains the same (Beran, 1994). Figure 14.5 shows an example of self-similarity in the daily attendance rates in one school (School 3). The first panel (Fig. 14.5a) shows the daily attendance rate in that school over a 7-year period, from the fall of 2004 through the spring of 2011. Figure 14.5b shows those rates for one school year (2007), and Fig. 14.5c shows the rates for the fall of 2007. Figure 14.5d shows the rates for a 22-day period within the fall of 2007. Comparison of these four trajectories suggests self-similarity in the following three ways. There appears to be a slight downward trend in Fig. 14.5a that replicates itself at the smaller grid levels of Fig. 14.5b, c, and d. In addition, there

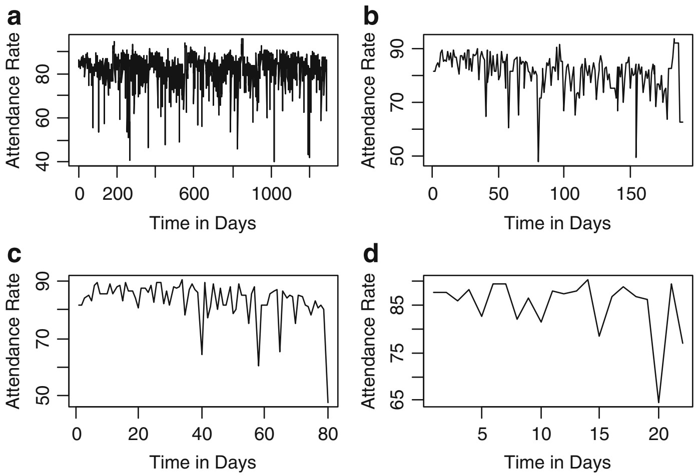

Fig. 14.5 Evidence of self-similarity in the attendance trajectory in School 3. (**a**) Attendance rates over a 7-year period (2004–2011, $N = 1290$); (**b**) rates in the same school over a 1-year period (2007, $N = 190$); (**c**) rates over the fall of 2007 only ($N = 80$); (**d**) rates over 22 days in the fall of that year

are pronounced dips that surface more toward the end of the series. Furthermore, at each level of description, variability appears to increase as the series progresses.

While this self-similar pattern is striking, not all features of the three trajectories replicate across scales: the last few observations toward the end of the trajectories show different variability patterns, and the lower dips do not necessarily occur at the same relative position of the time window. In the face of these conflicting signs, further statistical modeling is needed to empirically confirm the impression that daily attendance trajectories are indeed self-similar. As with the estimation of sensitivity to initial conditions, a large number of data points is needed to estimate a process hypothesized to replicate itself over and over in a scale-invariant manner.

The conventional ARIMA model described above is highly suitable to estimate such short-range dependencies, and a successfully fitted ARIMA model results in randomly distributed residuals. However, ARIMA models are not well suited for the detection and estimation of long-memory effects. The ARFIMA model is specifically designed to analyze the long-term fractional process that indicates self-similarity, by estimating the significance of the parameter d (the differencing parameter) over and above that of the autoregressive and moving average parameters. The use of ARFIMA to estimate long-range processes presumes a stationary trajectory, however. In case of non-stationarity, the investigator has the choice of analyzing the first (or second) difference, of the series, or resorting to different estimation methods altogether to detect fractality (Stadnitski, 2012a).

You may recall that in the short-term ARIMA (p, d, q) model, the parameter d is fixed to be zero if the trajectory is stationary, or $d = 1$ if it is non-stationary, in which case the first difference $Y_t - Y_{t-1}$ is analyzed. The *fractional* part of the ARFIMA (p, d, q) process refers to the fact that the detection of self-similarity through modeling of the long-range processes involves estimating fractions of d falling between $d = 0$ and $d = 1$. Dealing with the stationary case, ARFIMA also presumes that the differencing parameter d ranges from -0.5 to 0.5, with a $d = 0$ indicating no error dependency (white noise). A positive differencing parameter indicates a long-range positive autocorrelation pattern, also known as *persistence*. Conversely, a negative differencing parameter indicates a long-range negative autocorrelation pattern, referred to as *anti-persistence* (Beran, 1994; Stadnitski, 2012a; 2012b). Appendix 2 explicates ARFIMA formally.

A simulation with 1200 data points is shown in Fig. 14.6a, b to illustrate, respectively, short-range dependency in a simulated autoregression ($\varphi = 0.5$) and long-range dependency in a simulated fractal process ($d = 0.35$). The panels on the left show the relative stability of the autoregressive process in Fig. 14.6a compared to the more turbulent manifestation in Fig. 14.6b. The panels at the right of the figure show the characteristics of the corresponding ACF plots. The spikes indicating the size of the autocorrelations quickly recede to non-significance as the lag size increases in Fig. 14.6a, while in the plot in Fig. 14.6b the recession to non-significance proceeds very slowly, indicating persistence.

Figure 14.7a shows the mean-centered attendance trajectory in School 2 (see Fig. 14.5a for the original trajectory for this school), as well as the ACF plots at

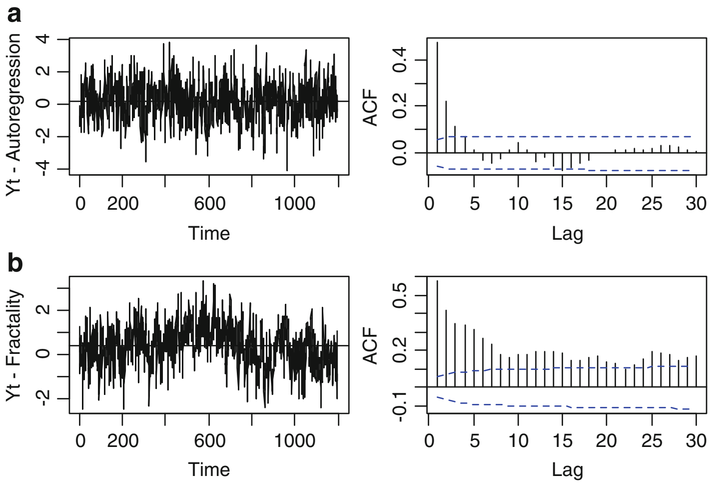

Fig. 14.6 Time series plot (*left* panels) and ACF plot (*right* panels) of (**a**) simulated autoregression ($N = 1200$, $d = 0$; $\varphi = 0.5$) and (**b**) simulated fractality ($N = 1200$, $d = 0.35$)

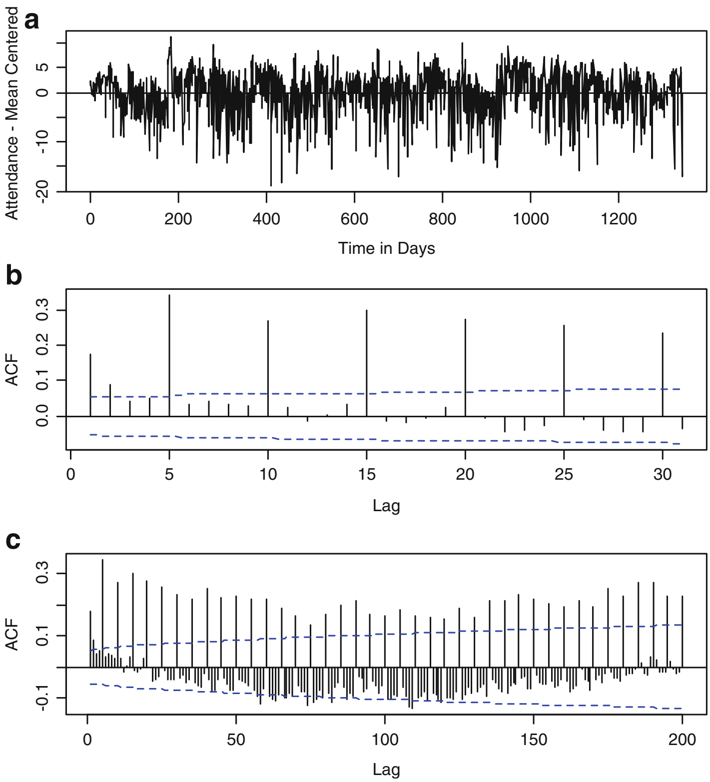

Fig. 14.7 Empirical data with short- and long-range dependencies (School 3); (**a**) attendance rates (mean centered); (**b**) ACF at 31 lags; (**c**) ACF at 200 lags

31 lags and at 200 lags (Fig. 14.7b and c, respectively). The short-term picture in Fig. 14.7b shows a rapidly decaying autocorrelations at the first few lags, as well as a seasonal cycle at the fifth lag that looks quite persistent. The longer term picture shown in Fig. 14.7c shows the persistence of the seasonal dependency as well as some evidence of nonseasonal persistence. Koopmans (2015) describes in greater detail the ARFIMA modeling process through which it was determined that the long-range dependencies made a statistically significant contribution to the variability in the data, even after modeling the short-range and seasonal processes in the trajectory. This analytical process yielded a differencing parameter of $d = 0.13$, indicating some degree of persistence over and above the short-range and seasonal dependencies.

In addition to the differencing parameter d, several other parameters are often used to characterize the dynamical process in time series data. One is the Hurst exponent H, named after Harold E. Hurst, who developed the measure to characterize the scaling dimension in such natural phenomena as water discharges, tree rings, temperature, and precipitation. Hurst originally defined the parameter in terms of the range R and standard deviation S of the measurement trajectories within given time periods to assess how the observed variability depends on the time ranges over which the measurements are taken. A linear correlation between the time range and measurement variability indicates long-range dependencies (Feder, 1988; Mandelbrot, 1997). Within the ARFIMA framework, the estimation of H is based on the differencing parameter d as $H = d + 0.5$. So then the interpretation of the differencing parameter provided above translates into an interpretation of the Hurst exponent as follows: $H = 0.5$ indicates white noise, $H > 0.5$ indicates persistence, and $H < 0.5$ indicates anti-persistence. Hence, the scaling component for School 3 equals $H = 0.63$, again indicating some persistence.

Power Spectral Density

To analyze fractal patterns in time series data, it is common practice to generate power spectra to assist with the detection of self-similarity. The conversion of a time series to a power spectrum involves a mathematical operation called Fourier transform (Shumway & Stoffer, 2011), which re-expresses the trajectory of observed measurements over time as a power versus frequency relationship as follows (Mandelbrot & van Ness, 1968):

$$S(f) \propto 1/f^{\beta}$$

In this function, f represents the frequency and $S(f)$ is the squared amplitude corresponding to that frequency (Delignières et al., 2005). The amplitude, or power, represents the magnitude of the variability in the cycles of dependency between observations at different lag values. The frequency in the spectrum is a *relative frequency*, which expresses the periodicity of the dependencies as $f = \frac{j}{n}$

with $j = 0, 1, 2, \ldots, (n-1)/2$. Here, j represents the number of cycles and n the number of time points in the series (Shumway & Stoffer, 2011). Thus, the relative frequency ranges from $\frac{1}{n}$ to $\frac{1}{2}$ after the Fourier transform is carried out. Few iterations j represent the long-term process and many iterations represent the short-term (Eke et al., 2000), and the power or amplitude expresses the strength of the dependency between the observations that constitute the cycle.

A power spectrum is produced by log-transforming the relative frequency as well as the power of this function. A power spectral density plot is said to display a *power law* if the relationship of the log power to the log relative frequency is linear with a negative slope. Such a relationship would indicate that if one, for instance, doubles the frequency, the power diminishes by the same rate regardless of the frequency values chosen on the abscissa of the density plot. This feature indicates the *scale invariance* that is one of the signature characteristics of self-similarity (Eke et al., 2000). Generating power spectra is therefore of theoretical as well as diagnostic interest in such cases. The parameter β in the power function above is used to estimate the slope in this plot, assuming that the relationship is linear. This latter proviso is an important reminder that a careful inspection of the power spectrum is required to determine whether this assumption is actually met. For a lucid discussion of the interpretation of linear and nonlinear patterns in power spectra, see Wagenmakers et al. (2004).

A major advantage of power spectral density analysis over ARFIMA is its capability of distinguishing fractality in stationary as well as non-stationary trajectories, typically referred to as fractal Gaussian noise (fGn) and fractal Brownian motion (fBm). As you may recall, ARFIMA requires stationarity in the data, and in the absence thereof, differencing is used to make the data stationary. Some researchers have argued that such a transformation effectively removes intrinsically interesting features from the data, resulting in information loss (Granger & Joyeux, 1980). Comparison of Fig. 14.2b and c illustrates this point. The differencing accomplished in Fig. 14.2c removes many interesting particularities from the data trajectory, such as the lack of consistency of the behavior of the data from one time period to the next in the original series. This feature, which could have major substantive interest in the analyses at hand, completely disappears in the differenced transformation shown in Fig. 14.2c.

The estimation of the Hurst coefficient H, on the basis of which the presence of long-term dependencies is decided, requires a distinction between fGn and fBm processes. Eke et al. (2000) describe how power spectral density analysis can be used to that end. The criteria for deciding whether a given time series belongs to the fBm or the fGn family are described as follows: if the slope of the power spectrum based on an observed time series equals $-1 < \hat{\beta} < 0.38$, fGn should be assumed when estimating H. If $1.04 < \hat{\beta} < 3$, fBm should be assumed. If $0.38 < \hat{\beta} < 1.04$, the process is said to be unclassifiable in terms of fGn vs. fBn. In the fGn case, the theoretical relationship between the Hurst exponent and the

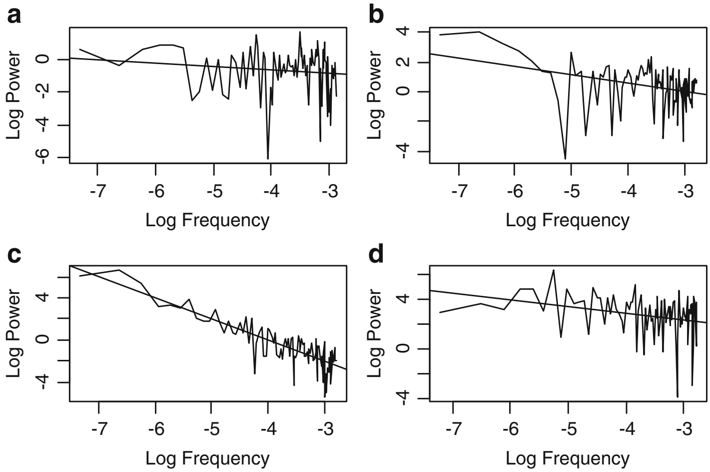

Fig. 14.8 Power spectra for (**a**) simulated white noise ($\beta = -0.21$), (**b**) simulated pink noise ($\beta = -0.57$), (**c**) simulated Brownian motion ($\beta = -2.0$), and (**d**) empirically observed daily attendance rates in School 3 ($\beta = -0.55$). The slopes for the simulated pink noise and Brownian motion spectra and for the attendance rates for School 3 are different from zero ($p < 0.05$)

power exponent β is $H = (\beta + 1)/2$; the power exponent can also be expressed as twice the differencing parameter estimated in ARFIMA, i.e., $\beta = 2d$. As indicated above, the Hurst exponent can be defined as $H = d + 0.5$. In the fBm case, $H = (\beta - 1)/2$. In both fGn and fBm processes, an H of 0.5 marks the boundary between persistence and anti-persistence (Stadnitski, 2012a, 2012b).

Figure 14.8a–c shows, respectively, what these power spectra would look like for simulated trajectories of white noise (no memory), pink noise (long-range memory), and Brownian motion (infinite memory). The ARFIMA simulation routine was used to generate white noise at $d = 0$ ($H = 0.5$) and pink noise at $d = 0.35$ ($H = 0.85$). Brownian motion was generated using phytools (Revell, 2012) with $\beta = 2.0$. In all three cases, the series were set to be 1500 observations long with a random normal distribution. The fourth panel (Fig. 14.8d) shows the power spectrum for School 3 ($N = 1290$ observations). The similarity between the power spectra for School 3 and the simulated pink noise in Fig. 14.8b is clearly discernible here, as are the differences between those two spectra on the one hand, and the white noise and Brownian motion spectra on the other. These differences can be appreciated both in terms of the steepness of the slopes and in terms of the amount of variability left around the fitted lines. As expected, the power function for white noise is flat; the slopes for the pink noise spectra fall well within Eke et al.' range for fGn, while the power spectrum for Brownian motion shows steeper slope with a narrow range of variances throughout indicating infinite memory (continued autocorrelation) in the entire trajectory.

Discussion

Few would argue that time plays a role in daily classroom and school-building activities, and the analysis of what the contingencies are that affect behavior in those contexts is a highly relevant undertaking as it might tell us what the underlying processes are of the transformations that constitute learning (Vygotsky, 1978). However, in educational science, our models of causal attribution tend to be cross-sectional, as we examine whether our instruction, leadership, policy, and other interventions impact the educational outcomes of our student population. These models are incomplete if the endogenous process is overlooked (Koopmans, 2014b). We need to know the behavior of interest over a larger time spectrum in order to understand the system's propensity toward transformation or toward maintaining the status quo. Knowing these propensities is important to qualify our causal attributions about educational effectiveness. Finding no relationship between interventions and outcomes may indicate that the system is resistant to change regardless of the (perceived) merits of the intervention in question. Likewise, it is possible that observed changes are not sustained in the long run in a highly flexible system as it deals with ever-changing adaptive requirements without sustaining the innovations whose effectiveness was demonstrated. We therefore need to acquire more knowledge about the internal systemic processes and how they behave over time because they reveal the system's predisposition toward change. Dynamical theories such as chaos theory and the theory of self-organized criticality are particularly concerned with such systemic propensities.

This chapter addresses two interrelated issues. The first one is that when the phenomena we study potentially have a temporal dimension, as may educational variables do, the contribution of this time dimension to the variability in one's observations needs to be investigated in a fair amount of detail to provide some meaningful answers about how endogenous processes contribute to the transformative process in education. Researchers may counter that longitudinal approaches such as survival analysis, repeated measures analysis of variance, and growth modeling can address this concern. However, these approaches differ from the ones described here in that traditional longitudinal techniques do not provide the degree of detail and resolution in the data that is required to estimate dynamical processes such as cyclical trends, or processes pointing to complexity such as self-organized criticality and sensitive dependence on initial conditions. The circumstances under which the time series approaches described are capable of capturing such complexity are a point of some contention in the dynamical literature. Eke et al. (2000) tested the reliability of fractality estimates using a time series of 2^{17} ($N = 131{,}072$), a length that is unlikely to have any meaningful empirical referents in education. Many researchers have proceeded with series of 2^9 or 2^{10} deemed sufficient for that purpose (Delignières et al., 2005; Stadnitski, 2012a). The challenge for nonlinear time series, in education as well as elsewhere, is the resource intensiveness of collecting information at this level of detail.

There is also a general point to be made about the cross-sectional use of central tendency and variability measures to address questions of educational effectiveness. The use of these measures presumes that the characteristics of interest are stable over time and that the time factor therefore does not have to be measured (the ergodic assumption, Molenaar, 2004). Given the dynamical nature of educational processes, it does not seem likely that the ergodic assumption holds very often; yet there are very few examples of the type fine-grained analyses that are needed to examine quantitatively the influence of time on the variability in our observations. This chapter illustrates one way of addressing this issue. Obviously, daily attendance rates are not the only variable of interest in the educational context. Important work to address the influence of time on educational outcomes and implementation variables includes several of the chapters included in this volume (Garner & Russell, Chap. 16; Pennings & Mainhard, Chap. 12; van Vondel, Steenbeek, van Dijk, & van Geert, Chap. 11), although the estimation of fractality is not the focus of that work.

The second concern addressed in this chapter is the fact that we know very little about how time contributes to school-level daily attendance rates in particular. The availability of a data repository covering more than a decade's worth of data by now has provided a unique opportunity to investigate the applicability of nonlinear time series in education, and learn more about how such attendance rates behave over the longer term. The analysis presented here indicates that in addition to the first-order autoregressive and moving average parameters that enhance the reliability of our descriptions, effective models incorporate seasonal estimators. Here, these estimators indicate that the 5-day weekly cycle exercises considerable influence over the patterns of variability found in the attendance trajectories. Practitioners may have been able to tell us about the seasonality of the daily attendance in their school buildings, but the formal research on high school attendance has traditionally had remarkably little to say about those patterns.

Of particular interest in the context of complexity research are the patterns in daily attendance that go over and above the seasonal influences noted above. The estimation of fractality or self-organized criticality is of interest because it points to complexity in the system as it adjusts to changing circumstances (Beran, 1994; Stadnitski, 2012b). This may be the case for schools as well, where schools showing fractality may have greater susceptibility to those influences whereas schools whose attendance trajectories do not show fractality may be more immune to them (Koopmans, 2015). Another aspect that is of importance to this discussion is the presence of many extreme observations that are likely to be tied to specific contingencies, such as snow days, upcoming vacations, and the irregularities associated with the end of the school year (Koopmans, 2011). Figures 14.1a, 14.4a, and 14.5a in this chapter illustrate the prominence of these observations. Irregularities of this kind are highly influential to the attendance trajectories, but they can usually be explained in terms of specific external contingencies, whereas the cyclical and long-range dependencies are often not as easy to account for. Particularly in those cases where schools show evidence of self-organized criticality or fractal patterns in the trajectories, the development of strong theories to explain those patterns becomes a pertinent issue, requiring investigator to collect additional information about putative causal influences such as parental support, teacher quality, and school responsiveness to student absences.

To develop strong causal theories about attendance behavior, in other words, it is necessary to triangulate the quantitative characteristics of school-level daily attendance trajectories with data from other sources to find out more about the factors that produce irregularity in attendance behavior as well as what the determinants are of self-organized criticality in the schools. Are small high schools to be more likely or less likely to display self-organized criticality? Are schools serving predominantly students from poor families more or less likely to show such patterns? It is up to empirical research to address these questions to help us understand better why given attendance rates are what they are, as well as to theory to articulate the putative causal mechanisms.

In this context, it is also relevant to contemplate what it means to say that there is self-organized criticality in the attendance trajectory for a given high school. Figure 14.5 in this chapter illustrates what it looks like in one school. The patterns shown there seem to suggest a fatigue dynamic, where initial cycles of high attendance/low variability are followed by higher variability and then lower peaks. This pattern appears to replicate in this trajectory in a scale-invariant manner, which is to say that it occurs over large time frames (e.g., a 7-year period), but also in much smaller time frames residing within those larger ones. The value of attendance research from a complexity perspective is that, contrary to the seasonal cycles that are easy to discern for school-building practitioners, these self-similar patterns are much harder to detect let along confirm, while they nonetheless have important implications for policy.

To estimate fractality, Wagenmakers et al. (2004) recommend a competitive modeling approach, along the lines of a stepwise multiple regression, where the statistical goodness-of-fit models including all the short-term estimators of interest are compared to a model including all of those as well as a differencing parameter estimate. Koopmans (2015) shows the applicability of this modeling strategy to daily school attendance trajectories. The literature advises caution when concluding self-organized criticality based on evidence of persistence in time series data, because the possibility remains that the appearance of persistence may in fact mimic a pattern of short-range dependencies (note that d is not lag specific in the general ARFIMA formulation shown in Appendix 2). Therefore, a careful inspection of the plotted trajectories and ACF plots is always indicated, as well as the triangulation of the statistical evidence from ARFIMA with other sources of information that may provide a more substantive description of the dynamics underlying long-range dependencies to help develop a strong causal theory to explain the results of time series analyses.

In closing, I'd like to stress the merits of single-case designs to enhance our understanding of educational processes, and the attendance data presented here are meant to illustrate that point as well. We can learn from studying the particularities of bounded individual systems and investigate in great detail the processes of self-maintenance and transformation as they play out over a large time spectrum and in the interactions between individual agents within the system (students, teachers, administrators, policy makers) and the larger systemic components (classrooms, school buildings, districts, federal agencies) with which these agents interact in an ongoing dynamical interrelationship. In educational science, a distinction is traditionally made between qualitative research, which tends to focus on the particular and quantitative research, which is oriented toward the analysis of data for purposes of statistical inference. The research presented here argues from a complexity angle for the obsolescence of the idea that quantitative and qualitative research are mutually exclusive empirical strategies. The richness of detail provided by the single case uniquely allows for a rigorous quantitative assessment of the dynamical underpinnings of behavior, as well as revealing its qualitative transformations.

Appendix 1: Short-Range Estimation Using ARIMA

The general model AR can be stated as

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

This model estimates Y_t using p lags. The parameter ϕ estimates the influence of past observations on the series at each given lag.

The MA model estimates Y_t in terms of accumulated error disturbances, also called innovations. Using q lags, this estimation can be written as follows:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

In this equation θ estimates the impact of each innovation on the series.

AR and MA processes can be captured in a single predictive model. For purposes of clarity, we describe a predictive model that uses one lag only, i.e., $p = 1$ and $q = 1$:

$$Y_t = \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1}$$

A special case is the *seasonal ARMA process*, which estimates the dependencies in terms of days of the week, months in a year, etc. The analysis presented here focuses on the regularities as a cyclical weekly pattern with 5 days in the school week. The model used to address this question can be formally written as

$$Y_t = \phi_1 Y_{t-5} + e_t - \theta_1 e_{t-5}$$

The autocorrelation function (ACF) at lag k is defined as

$$r_k = \sum_{t=k+1}^{n} \left(Y_t - \overline{Y}\right)\left(Y_{t-k} - \overline{Y}\right) / \sum_{t=1}^{n}\left(Y_t - \overline{Y}\right) \quad \text{for } k = 1, 2, \ldots$$

Appendix 2: Long-Range Estimation Using ARFIMA

Some mathematical reorganization of the terms in the ARIMA model as stated in Appendix 1 is required to describe what the estimation of the long-range influences adds to the models that assess the short-range effects on attendance trajectories.

It is often conventional in time series notation to express ARMA processes in terms of the so-called *lag operator*, or *backshift operator*, which is defined as

$$BY_t = Y_{t-1}$$

In plain English, the backshift operator B shifts observations back one time unit to construct a new series. The next lag over can be written as $BBY_t = Y_{t-2}$, or

$$B^2 Y_t = Y_{t-2}$$

In terms of this operator, the ARIMA process described above is often written as

$$\left(1 + \varphi_1 B + \varphi_2 B^2 + \cdots + \varphi_p B^p\right) Y_t = \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) e_t$$

The left side of the equation represents the autoregression (AR) component; the moving average (MA) component is on the right. The mathematical derivation of this formulation, called the *characteristic equation*, from the equations above can be found in Box and Jenkins (1970), Cryer and Chan (2008), and many other standard time series texts. It is assumed in this model that remaining error is randomly distributed, i.e.,

$$e_t \; (t = 1, \; 2, \; \ldots) \sim N\left(0, \; \sigma^2\right) \text{ IID.}$$

The ARFIMA model separates long-term dependencies from the short-term ones by parameterizing d as a differencing estimate:

$$(1 + \varphi_1 B + \varphi_2 B^2 + \cdots + \varphi_p B^p)(1-B)^d Y_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) e_t$$

It is assumed here that the trajectory is stationary and that $-0.5 < d < 0.5$ (Beran, 1994; Sowell, 1992; Stadnitski, 2012b).

References

Ashby, W. R. (1957). *An introduction to cybernetics*. London: Chapman & Hall.

Astone, N. M., & McLanahan, S. S. (1991). Family structure, parental practices and high school completion. *American Sociological Review, 56*, 309–320.

Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York: Springer.

Balfanz, R., & Byrnes, V. (2012). *The importance of being in school: A report on absenteeism in the nation's public schools*. Baltimore, MD: Johns Hopkins University Center for Social Organization of Schools.

Beran, J. (1994). *Statistics for long-memory processes*. New York: Chapman & Hall.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis, forecasting and control*. San Francisco, CA: Holden-Day.

Cryer, J., & Chan, C. (2008). *Time series analysis with applications in R*. New York: Springer.

Delignières, D., Torre, K., & Lemoine, L. (2005). Methodological issues in the application of monofractal analyses in psychological and behavioral research. *Nonlinear Dynamics, Psychology, and Life Sciences, 9*, 435–462.

Eke, A., Herman, P., Bassingthwaighte, J. B., Raymond, G. M., Percival, D. B., Cannon, M., et al. (2000). Physiological time series: Distinguishing fractal noises from motions. *Pflügers Archives, 439*, 403–415.

Feder, J. (1988). *Fractals*. New York: Plenum.

Fuller, W. A. (1996). *Introduction to statistical time series* (2nd ed.). New York: Wiley.

Glass, G. V. (1972). Estimating the effects of interventions into a non-stationary time series. *American Educational Research Journal, 9*, 463–477.

Goldstein, J. (1988). A far-from-equilibrium systems approach to resistance to change. *Organizational Dynamics, 17*, 16–26.

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-range time series models and fractional differencing. *Journal of Time Series Analysis, 1*, 15–30.

Hamilton, P., Pollock, J. E., Mitchell, D. A., Vincenzi, A. E., & West, B. J. (1997). The application of nonlinear dynamics to nursing research. *Nonlinear Dynamics, Psychology, and Life Sciences, 1*, 237–261.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika, 68*, 165–176.

Hurst, H. E. (1965). *Long-term storage: An experimental study*. London: Constable.

Jensen, H. J. (1998). *Self-organized criticality: Emergent complex behavior in physical and biological systems*. Cambridge: Cambridge University Press.

Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* (2nd ed.). New York: Cambridge University Press.

Kaplan, D., & Glass, L. (1995). *Understanding nonlinear dynamics*. New York: Springer.

Kemple, J. J., Segeritz, M. D., & Stephenson, N. (2013). Building on-track indicators for high school graduation and college readiness: Evidence from New York City. *Journal of Education for Students Placed at Risk, 18*, 7–28.

Kemple, J. J., & Snipes, J. C. (2000). *Career academies: Impacts on students' engagement and performance in high school*. New York: MDRC.

Kerlinger, F. N. (1977). *Foundations of behavioral research* (2nd ed.). London: Holt, Rinehard & Winston.

Koopmans, M. (2009). Epilogue: Psychology at the edge of chaos. In S. J. Guastello, M. Koopmans, & D. Pincus (Eds.), *Chaos and complexity in psychology: The theory of nonlinear dynamical systems* (pp. 506–526). New York: Cambridge University Press.

Koopmans, M. (2011, April). *Time series in education: The analysis of daily attendance in two high schools*. Paper presented at the annual convention of the American Educational Research Association, New Orleans, LA.

Koopmans, M. (2014a). Nonlinear change and the black box problem in educational research. *Nonlinear Dynamics, Psychology and Life Sciences, 18*, 5–22.
Koopmans, M. (2014b). Change, self-organization, and the search for causality in educational research and practice. *Complicity: An International Journal of Complexity in Education, 11*, 20–39.
Koopmans, M. (2015). A dynamical view of high school attendance: An assessment of the short-term and long-term dependencies in five urban schools. *Nonlinear Dynamics, Psychology and Life Sciences, 19*, 65–80.
Lewin, K. (1947). Frontiers in group dynamics. *Human Relations, 1*, 5–41.
Mandelbrot, B. B., & van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review, 10*, 422–437.
Mandelbrot, B. B. (1997). *Fractals and scaling in finance: Discontinuity, concentration, risk*. New York, NY: Springer.
Molenaar, P. C. M. (2004). A manifesto on psychology as an idiographic science: Bringing the person back in to scientific psychology, this time forever. *Measurement, 2*, 201–218.
National Center of Education Statistics. (2008). *Average daily attendance in public elementary and secondary schools, by state or jurisdiction: Selected years, 1969–70 through 2005–06*. Retrieved June 30, 2010, from http://nces.ed.gov/programs/digest/d08/tables/dt08_040.asp
Neter, J., Wasserman, W., & Kutner, M. H. (1985). *Applied linear statistical models: Regression, analysis of variance, and experimental designs*. Homewood, IL: Irwin.
Peng, C. K., Mietus, J., Hausdorff, J. M., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1993). Long-range anti-correlations and non-Gaussian behavior of the heartbeat. *Physical Review Letters, 70*, 1343–1346.
Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution, 3*, 217–223.
Roby, D. E. (2003). Research on school attendance and student achievement: A study of Ohio schools. *Educational Research Quarterly, 28*, 4–15.
Shumway, R. H., & Stoffer, D. S. (2011). *Time series and its applications* (3rd ed.). New York: Springer.
Sowell, F. (1992). Modeling long-run behavior with the fractional ARFIMA model. *Journal of Monetary Economics, 29*, 277–302.
Sprott, J. C. (2003). *Chaos and time series analysis*. New York: Oxford University Press.
Stadnitski, T. (2012a). Some critical aspects of fractality research. *Nonlinear Dynamics, Psychology, and Life Sciences, 16*, 137–158.
Stadnitski, T. (2012b). Measuring fractality. *Frontiers in Physiology, 3*, 1–13.
Sulis, W. H. (2012). Causal tapestries for psychology and physics. *Nonlinear Dynamics, Psychology and Life Sciences, 16*, 113–136.
Thornton, P. L., & Gilden, D. L. (2005). Provenance of correlations in psychological data. *Psychonomic Bulletin & Review, 12*, 409–441.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Wagenmakers, E. J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of $1/f^{\alpha}$ noise in human cognition. *Psychonomic Bulletin & Review, 11*, 579–615.
Waldrop, M. M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.
Wiener, N. (1961). *Cybernetics, or control and communication in the animal and the machine* (2nd ed.). Cambridge, MA: MIT Press.