## The Science of Forecasting

Forecasting is what most crime analysts want from temporal analysis. Many are also intimidated by it—after all, as analysts we stake our credibility on each forecast. Many agencies or bureaus within agencies even have policies restricting forecasting. Should this be the case? Forecasting is definitely risky—nobody can read the future, and as professionals we shouldn't pretend that we can. But what's the point of analyzing crime if we never actually do anything about it?

The basic assumption that underlies all forecasting is that what will happen in the future will resemble what has happened in the past.

The term "forecasting" when used in reference to tactical crime analysis specifically describes the process of generating a prediction as to where, when, and how future events in a crime trend, pattern, or series will occur. Forecasts are generated by studying known events in the series to identify patterns that might prove susceptible to prediction. Most crime analysts who have performed any tactical analysis in their careers have made some attempt to forecast a crime trend, pattern, or series.

Strategic analysts also make forecasts. Their efforts, however, are almost entirely statistical in nature, and relate to the "big picture" view of crime and policing within the scope of their observations and analyses. It is perfectly in keeping with the nature of strategic forecasting to predict, for example, "There will be an 11 percent increase in sexual assaults over the next three months." This prediction can provide law enforcement agencies with broad intelligence on what their future efforts will entail, enabling them to prepare in advance by altering their patrol plans, investigative priorities and schedules, public awareness campaigns, and so on.

A tactical forecast, however, has as its goal the prediction of a specific crime, usually by a specific perpetrator. This is a vastly more complicated affair. Why so? The difficulty has to do with the nature of linear versus chaotic systems.

For thousands of years, scientists have attempted to describe the world in terms of "linear system" models. Linear systems follow clear-cut rules and are perfectly predictable. For example, the movement of stars and planets was early discovered to be precisely predictable, and is an example of a linear system. The movement of a pendulum is also an example of what is considered a linear system: with the right equations, one can precisely predict the movement of any pendulum. The trajectory of a ballistic projectile is a third illustrative example. If one knows the angle, inclination, mass, and force behind, say, an artillery shell, it is possible to predict precisely where it will fall.

Recently, however, scientists have begun to describe the world in terms of "chaotic system" models. Chaotic systems, unlike the clear and simple linear models, are very complex, and must take into account the interaction of an enormous number of variables. One of the most often-cited examples of a chaotic system is weather. It is impossible to predict the weather accurately, because so many variables intertwine in unpredictable ways that there is no mechanism for embracing them all. In fact, one accepted tenet dealing with the nature of subatomic particles (the "Heisenberg Uncertainty Principle") states that the specific characteristics of such particles cannot be measured without changing them, and therefore they cannot be known; this means that many phenomena are utterly impervious to prediction using linear models.

In fact, every one of the linear systems cited as an example above is now usually considered to be chaotic; in many ways this is much more realistic approach. For instance, consider the movement of a pendulum. A physics instructor will tell you that the movements of a pendulum are completely predictable, based on the formula $T = 2(L/g)^{1/2}$. While this is true of a "perfect" pendulum, which exists only in our minds as a mathematical illustration, it is demonstrably not true of a real pendulum. This is because a real pendulum will contain minute imperfections, resulting in slight distortions in its balance; it will be subject to microscopic influences from humidity and air pressure, and from the immeasurably tiny seismic disruptions from people walking nearby, trucks moving on the roads a mile away, and so forth. Even the effects of mere photons, shining in through its glass case (and themselves unpredictably polarized and distorted by the glass and the atmosphere) will impart their miniscule force, disturbing the perfection of our mathematical model. But the physicist need not despair totally: After all, the influence of these tiny forces is often too small to be measured at all. So, even if the prediction is wrong, perhaps it's only inaccurate by so tiny an amount that we don't notice it at all.

This is exactly the sort of abstract, arcane, math-oriented ground that makes most professional analysts wince. How is this relevant? Let's put our feet back on the ground for a moment, and take another look at forecasting tactical crime events. Is crime a linear or chaotic system? The answer, which we hope is obvious, is that it is most definitely chaotic. In fact, although at first it might sound absurd, the relationship between the movements of subatomic particles and the actions of an individual makes a reliable metaphor. Just as the movements of a pendulum are not perfectly predictable due to the large number of individually tiny (but cumulatively significant) influences of subatomic particles, yet remain almost predictable, at least as far as our imperfect instruments can measure, so too the actions of many people reacting to the stimuli of our complex social environment are unpredictable—but we can come close. The more broadly we measure, the more reliable our predictions.

Why can't we make a perfect tactical forecast? Too many imponderable factors, running down from the cosmic to the microscopic, can influence the behavior of the offender and his targets. We might formulate a reasonable forecast, yet miss our mark due to the offender coming down with the flu, or getting a flat tire. Perhaps the business he planned to rob closed early for the holidays, or burned down in an unrelated accident. Perhaps the offender's mother is staying with him over the weekend, and he can't very well sneak away to rape someone unnoticed. There are infinite reasons why no prediction, however artful, can ever be absolutely reliable and accurate.

Given these seemingly insurmountable difficulties, then, one might be tempted to wonder why we should try at all? The answer is that we might still come close. Remember how we noted earlier that it is impossible to predict the weather? That's true; however, we can come pretty close, if we're careful. Meteorologists can make informed guesses as to when and where it will rain or snow, and, with the current state of technology and methodology, they're right far more often than they're wrong, in the short-term, at least. Whether it will rain two months from now is anybody's guess; but whether it will rain three days from now is a matter where a trained meteorologist with access to reliable data can make an informed prediction. Forecasting the future of a crime series involves predicting the behavior of an offender, relative to his victim, in space and in time. If this behavior were purely a matter of random chance, we would have poor prospects for success indeed! Offender behavior, however, is not random. Offenders make decisions to commit their acts based on their needs and desires, based on their own prejudices, knowledge of their environment, and assessment of the risk to themselves. They base their actions on their own sense of morality (or more properly lack thereof). Although we might seldom call these actions "reasonable" (Who reasonably commits murder? Or rape?), we can still consider them "rational." That is, the crimes are committed for a purpose, and according to a plan. The purpose may be utterly mysterious, even to the offender; and the plan may be purely subconscious, and appear arbitrary and slipshod—but nevertheless the offender is surely influenced by them. By carefully studying what we know of his behavior from past crimes, by analyzing how he seems to move in and perceive time and space, and by identifying his pool of potential victims, we stand a small but very real chance of extrapolating what his future actions are likely to be, even when he himself does not know, consciously.

This tells us what we can safely expect from tactical forecasts. Although we can never hope to be perfect, no matter how powerful our computers or elaborate our methods, we can definitely hope to come close enough to actually intercept a crime—the Holy Grail of tactical crime analysis. Many practitioners have enjoyed success in predictions; some have even refined their process into a repeatable method.

There are three main families of temporal forecasting: Linear, Cyclical, and Leading Indicator. **Linear** forecasts extrapolate what will happen based on an overall trend—a line—that is drawn from the past, through the present, and then carried on into the future. This is most easy to imagine in the context of linear visualizations of time: Timelines and tempograms. In linear temporal visualizations, the past and the future are in opposite directions, the farther you go in either direction, the more distant everything on the other end becomes.

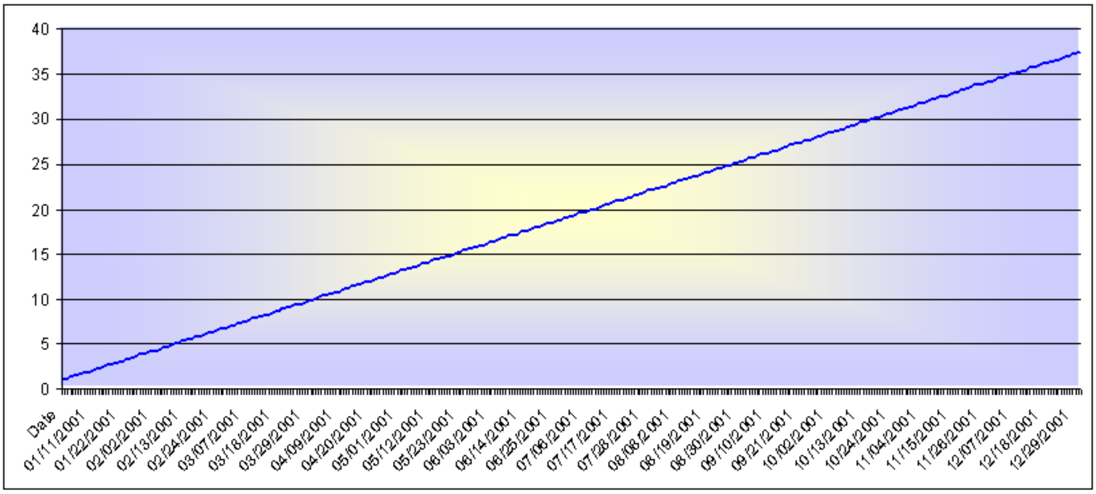


*Figure 11-38: A linear timeline*

**Cyclical** forecasts wrap around themselves—they make use of the circular, repeating patterns that we use to measure time. Things like seasons, months, days of the week, hours of the day—these things are extremely important to human behavior and, so, to criminal behavior too. But unlike a linear timeline or a tempogram of events, cyclical visualizations wrap around themselves. So, if we begin our week on Sunday, then proceed through the days of the week, we find, seven days later, we're right back where we started.

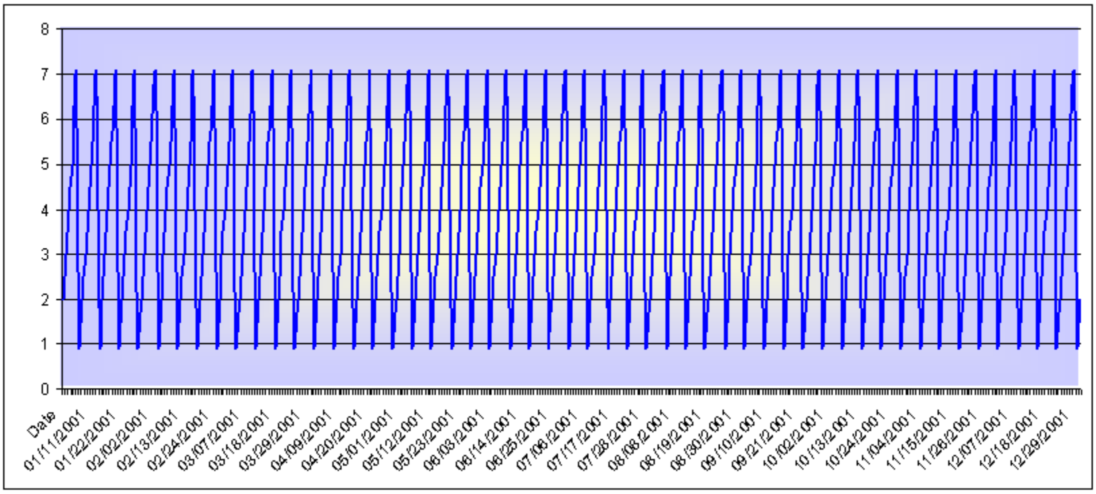


*Figure 11-39: A cyclical timeline*

These two main families give rise to a hybrid, more challenging family of temporal forecasting: complex forecasting. This type of forecasting incorporates both, by taking a linear trend and then either adding or multiplying cyclical values:

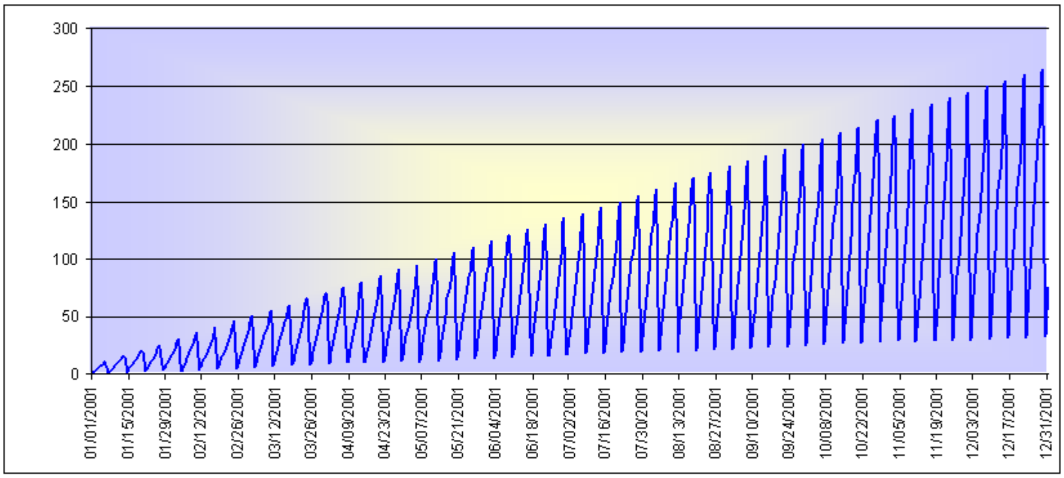


*Figure 11-40: A hybrid timeline, with both linear and cyclical characteristics*

The most commonly used model for performing complex temporal forecasting is known as ARIMA. This is an acronym for Auto-Regressive Integrated Moving Average, a process developed by the U.S. Census Bureau for population forecasting in the twentieth century. Several variations of the ARIMA algorithm have been employed for decades by a wide range of scientists and scholars. The X-12 variant is the current favorite; free utilities to implement this algorithm can be obtained from the U.S. Census Bureau.

**Leading Indicator** forecasting is bi- or multivariate in nature. Remember how we observed earlier that certain temporal behavior—like the timing of crimes—could be influenced by other variables—like the amount of money taken? Well, in that example, the amount of money taken would be a leading indicator—a variable that independently predicts the outcome of a dependent variable—timing. Financiers, demographers, and climatologists have long mastered the science of forecasting based on leading indicators. It can be surprisingly easy. As we've seen, bi- or multivariate analysis of our timeline might identify for us what variables may lead the timing of our series.

In temporal analysis, there are three things we can forecast: timing, frequency, and duration:

**Timing** forecasts attempt to determine when an event will take place. Tactical analysts use these forecasts to try to predict when the next event in a crime series will occur.

**Frequency** forecasts attempt to determine how many events will occur within a certain period of time. Strategic analysts use these forecasts to try to predict how many events will happen during some point in the future, such as how many robberies we will have next year.

**Duration** forecasts attempt to determine how long an event will last—we don't see these much, as crime analysts; however, administrative and operational analysts might attempt to determine how long officers are likely to spend on certain types of calls, how long it might take victims to report certain crimes, or how long it might take to peaceably disperse a crowd.

Timing forecasts are the particular challenge of the tactical crime analyst. A great many things have to go right for this type of forecast to succeed—but succeed it can, with thought.

How can we forecast when a crime will occur? At first approximation, that might seem as difficult as guessing the winning lottery numbers—but it's not that difficult, because the criminal's timing isn't random, and we can therefore study, model, and predict it. On the other hand, some analysts have forwarded simplistic forecasting methods, suggesting that something as simple as the mean interval method or hourly frequency forecasting can predict when a crime will occur—and indeed they can, sometimes. But, then, so can just guessing. The truth is that forecasting is neither impossible nor easy; it requires some careful consideration. When it comes to making temporal next-event forecasts, we have several options:

### *Linear Timing Forecasts*

Based on the interval between cases, we can guess what the probable interval from the last known case to the next (future) case will be. The mean interval method satisfies this requirement, but only under very restricted circumstances. A tempogram approach will solve all problems that would work with the mean interval method, and most other types of series, too, including both accelerating and decelerating crime tempos. This is the preferred method for performing linear timing forecasts. We can accomplish tempogram forecasting using an extrapolated trend line based on previous intervals. The trend line can be either an ordinary least squares (OLS) trend line, or a variation, such as the logarithmic least squares, which curves to approach, but never exceed, terminal velocity.