

Perils and Pitfalls of Empirical Forecasting

Peter Stallinga,
Igor Khmelinskii, (PhDs)
University of The Algarve, Portugal

doi: 10.19044/esj.2017.v13n18p18 [URL:http://dx.doi.org/10.19044/esj.2017.v13n18p18](http://dx.doi.org/10.19044/esj.2017.v13n18p18)

Abstract

It is common to use past information about the system modeled in probabilistic statistics to make predictions about the future. Especially in the area of climate modeling and forecasting this is done. Here it is argued that doing this in a purely empirical way is full of perils and pitfalls. Without knowledge of the underlying physical laws it will go wrong sooner or later. Specifically, the distribution functions are analyzed, which are normally assumed to be well-behaved gaussian-like not because there is a reason for it, but only because they don't cause mathematical problems. Real functions (like power laws) will prohibit any statistical analysis and thus prediction model. Furthermore, correlations and extrapolations are considered. The first show that correlations come in many types and not all of them have a direct causation link. The specific case of extreme events is used as an example to highlight the difficulty and the pitfalls of empirical forecasting in general. The conclusion is that empirical forecasting cannot be used for science.

Keywords: Empirical forecasting, extreme events, bell curve, extrapolation, correlations

Introduction

It has always been popular to analyze past events and make some kind of 'statistical analysis' of them to allow for making statements about the future of the system under study. This, actually, is common practice in many areas, such as the stock markets, the climate, finance, epidemics, etc. Society needs predictions as a base for implementing measures. The current handling of the economic crisis is a good example, where past financial data is used to do stress-testing of financial institutions. Yet, it is highly dubious whether this kind of approaches is adequate, even if it satisfies our inherent need to 'do something'.

The stock market is a nice example where we can recognize the underlying processes. We understand that the forecast itself can and is often influencing the outcome. This immediately proves, according to the efficient

market hypothesis (EMH), that the open and efficient market prices are not forecastable, since otherwise anyone would be able to attain unlimited gains (Timmermann and Granger, 2004); If we know for sure that the stock of X tomorrow will be higher than today by $y\%$, we'll buy already today and the stock will already be $y\%$ higher today and not tomorrow. A typical adage "Go away in May, but remember to return in September" is thus void and useless. It is at best market manipulation (the manipulator sells in April and comes back in August), and forecasting – stock recommendations – should be ignored by wise investors, since at best they are self-fulfilling prophecies that benefit only the forecasters. This is a commonplace, though.

But even in systems without feedback to the prediction itself (like natural processes), care has to be taken to not over-rely on the predictions. Purely empirical predictions, by definition, are based on past events only, without understanding the dynamics and underlying laws that govern the system. This paper discusses in a mostly philosophical way (of Aristotle) why a purely empirical forecasting (EF) approach is meaningless. We use the weather/climate forecasting to exemplify our points, because it is prototypical; it can serve as an archetype of a system where the empirical approach is widely used. The weather prediction is based on extremely poor knowledge of the system – atmosphere, oceansphere, landmasses, earth mantle, etc. As an example, so far nobody can predict El Niños and La Niñas (generally ENSO, El Niño Southern Oscillation) reliably, while the impact of them on the climate is large and irrefutable. Many other systems are in this category, no or poor knowledge of the laws combined with an abundance of data and thus an overuse of the empirical approach.

To give a (classical) example of pitfalls in EF, think of a roulette table. Imagine that the outcome was 'red' twenty consecutive times. There are now two extreme ways of looking at it, both considered 'cognitive biases' (wrong estimations of probabilities). We can either say that next time there is a high probability of winding up 'red' again ($\gg 0.5$), since not knowing anything about the table, obviously this table has a preference for 'red' as proven by past data. This is called 'hindsight bias' and is the purely empirical approach. The other bias is exactly the opposite, "It must be 'black' next time, since I know the physical processes and know it must be 50-50 over time and thus it must be 'black' now". This is called the Gambler's Fallacy. Both are wrong (for a non-fixed table). We focus here on the former bias, the purely empirical approach.

The idea of empirical forecasting even became a science in 1693 when Edmond Halley published his article on life annuities. By looking at the death tables of Breslau (now Wrocław), without any attempt of explaining these data, he calculated how much people in a community should set aside for allowing them to have a pension and receive a yearly 'annuity'

until death. A research area was born, Actuary, the mother of all empirical forecasting. By looking at past data – 'tables' – probabilities are determined and the future is planned. The (economical) success of this science ensured a wide dissemination of the ideas. In 2017 most events are seen as probabilistic whose distribution parameters can be found in past data. This ranges from the weather to the stock markets to sports events, etc.

With the advent of computational power, empirical forecasting got a new dimension. It now also includes 'artificial intelligence' approaches, like using Neural Networks (NN) or Principal Component Analysis (PCA) for finding correlations in and between data. However, there is no logical difference between human empirical analysis and computer-aided empirical analysis, although AI can help us do it faster.

A modern variant of the empirical approach that also involves large computational power is finite-elements calculations as for instance used in weather and climate predictions, and is a slightly improved version of the pure empirical approach in that it does have a small amount of knowledge in the system incorporated. In this method, the system is divided into many individually-acting tiny cells that interact with each other through well known laws (so seemingly not empirical). For instance, for the weather and climate the atmosphere is divided into millions of cells that each follow basic well-known physical laws, like classical gas laws, thermodynamics and fluid dynamics (Lynch, 2008). Similarly, in an artificial stock market approach, a stock market can be divided into many heterogeneous interacting adaptive traders, a system that thus shows emergent properties not seen at the level of individual agents (Chen and Yeh, 2002). (Interestingly, a similar analysis, when applied to interacting scientists, governed by the 'laws of publication', results in a rapid formation of a consensus among scientists on any subject that has high relevance to society, without the need for the consensus to represent the truth (Stallinga and Khmelinskii, 2014)).

The problems with this empirical approach are myriad. Starting with the least severe and increasing severity: i) The underlying laws may be incorrect, ii) They may not be applicable to the system under study (for example – very popular – applying physical laws like those of Statistical Physics to non-physical systems like the financial markets (Cordier, Pareschi, and Toscani, 2005)), iii) It is not clear if it is the only set of laws, iv) It is not clear if they are sufficient, v) boundary conditions or initial conditions may be unknown, i.e., the 'data' may be unreliable, vi) It is not clear if this reductionist approach (system is sum of parts) is adequate in the first place. vii) After successfully describing (retro-predicting) past data the empirical models might not work for future data, since the system may have changed, something that was not included in the simulation because it was not needed to explain the data. An example is the problem of obesity in

society, a health risk for the young people overrepresented in the current demography, but maybe not so for the abundant elder in the future (Gibbs, 2005). viii) Afterwards, in turn, the simulations have to be studied, which are as complex as the original system and nothing was gained in knowledge of the system. We are as far away from home as where we started. An example in Biology of this last point is the prediction of the tertiary and quaternary (3D) structure of proteins based on their primary (chemical) and secondary (grouping) structure (Holbrook, Muskal, and Kim, 1993); (Angiulli, Fionda, and Rombo, 2007). In fact, the best method, with 100% accuracy, is a black box in which are joined together a protein-construction machine and an X-ray analysis machine, both technologies already existing. It can 100%-accurately 'predict' the structure of a protein, but no knowledge whatsoever will be gained by doing so.

Sometimes the mere fact that the simulations can simulate the past is seen as a proof that the model is correct. Yet, this is dubious if no knowledge is gained. To show this, ad absurdum, imagine a computer program that retrodicts a time series $x_i = \{3, 5, 8, 3, 4 \dots\}$ that consists of the following code:

```
if (i=1) then x := 3
else if (i=2) then x := 5
...
```

which is obviously 100% correct, but has no underlying knowledge of the system and can thus not be trusted when used for predicting the future. In other words, the fact alone that the simulations can predict the past does not mean anything in itself. Imagine now adding after 100 real measurement data, explained with the above program, a prediction of the program

```
else if (i=101) then x := 200
```

We will believe our predictions for the data point (year?) 101 because our model manages to explain the past with complete accuracy. Who needs more proof of our model?

Philosophically, the empirical approach is a form of pure inductivism (induction plus reduction) which has been rejected as a correct approach since it is only applicable to past data (Chalmers, 1982); If you have seen only lemons in your life, you'll conclude that all citrus-fruit are yellow – no finite elements simulation will make you see it otherwise, because you'll adjust the parameters of your simulations until the outcome coincides with the data, that is, all citrus-fruit are yellow – and see your model shattered the first moment you see an orange. This is what Taleb calls a Black Swan (Taleb, 2011) and is the most severe pitfall of the empirical approach and inductivism in general.

Yet, scientific literature is full of papers on understanding past data. In an example of a study of the Forex market (dollar/rial), it was concluded

that “. . . profitable [. . .] trades could be made using past data” (Abounoori, Shahrazi, and Raseki, 2012). Of course, on past data profit could have been made, but will the algorithm also enable profit in the future? We suspect that the authors are not as rich as they were implying they could become. Another example is the analysis of a soccer championship or sports in general (Karlis and Ntzoufras, 1998); (Karlis and Ntzoufras, 2003). After the final match, it can be analyzed why exactly the winning team has won the championship. Physical strength, agility, endurance, intelligence, etc. The championship can be fully parametrized and retro-predicted. Yet, since a lot of the game is pure luck, it is highly unlikely the model will adequately predict next season’s result. The parameters will have seemingly changed next year. Suddenly, having a lot of Swedish players on your team will prove to be the decisive factor, while last year they had to be bearded, the latter a-posteriori explained by a sophisticated physical local-friction-coefficient argument (to be published in a fancy journal). Next year, the model will be adjusted in a Bayesian way and will again accurately retrodict the past. An example is the climate, previously unforeseen recent cooling fully retrodicted again (Guemas et al., 2013). Real data not used to reject the hypotheses, but used to adjust them instead. Another example of Bayesian ‘science’ – adaptive predictions – is summarized by Anthony Watts in his entry “We’ll have an ice free Arctic by the year xxxx” on his blog (Watts, 2013).

Why do we still spend so much time on these kinds of empirical analysis then, so many researchers doing this type of work, if it is so unreliable? The answer is quite simple. It is better than nothing, and maybe by sheer coincidence we will find some answers one day. Inductivism is part of the scientific method (Chalmers, 1982). There are many such papers in literature. That is also because it is easy to get such an analysis accepted in a journal; After all, it is very difficult to reject a manuscript that presents an idea that is consistent with the data. Yet, in the meantime, we should not heed so much the predictions and not give them the benefit of the doubt; “They are right because they are consistent with the data”. A winning theory will have been submitted to heavy scrutiny (falsification), and, among other things, shown to be the only theory that can explain the data and on top, be able to predict future data (without having to adjust the theory). As such, we actually need as many ideas and theories as possible. The fancier the better. Let people be creative and imaginative in their thinking. The erroneous papers will eventually be forgotten, and the few good ideas – those that made accurate predictions over and over again – will float to the top. In the next sections we will look at the probability functions used, discuss the difference between stochastic and deterministic processes, treat interpolation and extrapolation, and describe correlations between and in data.

Empirical statistics, the bell curve

The reader by now already has an idea of what we are trying to say from what was written above. The perils of empirical forecasting, which is in fact an implementation of what is in psychology called 'hindsight bias' (predicting the past, or 'retrodiction'). In this part we analyze the actual distribution functions used for empirical forecasting. These are normally of bell-curve types, or similar functions. Taleb, fully understanding the philosophical problem of the empirical forecasting, mentions them in his book *The Black Swan* by "Deluded by the simplicity of the bell curve" (Taleb, 2011). This we find a quite adequate description. We give here an example of weather and climate, since it is easiest understood. For this we first have to introduce the weather and the climate, the bell curve and then show where it goes wrong.

The World Meteorological Organization defines climate as the 30-year unweighted average weather normal that is to be recomputed every 10 years (Arguez and Vose, 2011), i.e., a semi-moving-averaging filter. A more human definition of weather and climate comes allegedly from Mark Twain: "Climate is what you expect, weather is what you get", which is nicely relevant for this work, since it includes both the past and future prediction parts of the subject. Thus, climate is the average of the past and is used as a prediction for the weather of the future. Nature, paying no attention to our human analysis, has its own agenda and basically ignores our predictions and gives us what we get. To give an example of contemporary climate reasoning, look at Figure 1. It shows the analysis of winter temperatures in Laputa (the country of Gulliver's Travels), similar to the plot used by the IPCC – the consensual authority on climate analysis – in their FAQ pages addressing the point if weather events are signs of climate change (IPCC, 2007). The figure shows that the data, until one exceptional year, followed nicely a Gaussian pattern; the probability density $p(x)$ of having a value x is given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

with μ and σ descriptive parameters, the first one being the expectation value and the second the spread. Synonyms for the 'Gaussian curve' are the 'bell curve' or 'normal distribution'; they are all identical. The bell curve was invented in the eighteenth century as a way to represent binomial probabilities, for example the fraction of heads in an infinite number of n coin tosses (Fendler and Muzaffar, 2008).

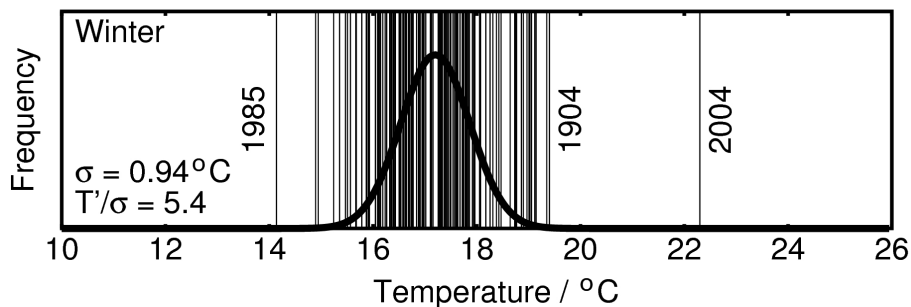


Fig. 1: Fictional data showing a normal distribution and an outlier of winter temperatures in Laputa (the country from Gulliver’s Travels). Similar to the plot of summer temperatures in Switzerland presented in FAQ 9.1 of Ref. (IPCC, 2007) by the alleged climate authority, the IPCC

If the past data falls nicely on a Gaussian distribution, this alone is seen as empirical proof that the distribution is Gaussian. (There is no other justification for the Gaussian assumption, other than the past data. Remember the roulette table). The reasoning is now that the exceptional data point is a sign of climate change, because it falls outside the prediction error margin, namely 5.4 ‘sigma’ (σ); The probability of finding a value so far away from the mean can be calculated with the cumulative probability function,

$$\begin{aligned}
 P(x \geq X) &= \int_x^{\infty} p(x)dx \\
 &= \frac{1}{2} [1 - \text{erf}(\frac{X-\mu}{\sigma\sqrt{2}})]
 \end{aligned}
 \tag{2}$$

and for $X = \mu + 5.4\sigma$ this is too tiny (0.0000033%) to be reasonably possible. It is more likely that the system (parameters) – the climate – has changed. This, at least, is the reasoning used by the climate reference institute, the IPCC. However, this is a form of circular reasoning, a logical pitfall. Experimental data is used and forced into a certain distribution (Gaussian here), and then any data not conforming the distribution statistics is called an ‘outlier’ and thrown away – as often done by ‘stickers to the model’; ”Trust your model, facts can be disposed of” – or used as proof of parameter changes (μ models the climate, i.e., the prediction of the weather; μ change is climate change and thus tomorrow’s weather change), as done by people who want to change the model in a Bayesian, adaptive way. (An example is Lovejoy who analyzed the bicentennial temperature data in a stochastic concept with Gaussian-like distributions and concludes with high level of confidence that the climate changes are man-made (Lovejoy, 2014)).

Nobody’ll say that the assumption of the (Gaussian) probability distribution is wrong in the first place, even while that might be the most

logical conclusion. Even worse, the authors of the altered model do not even realize that the outlying data point is also inconsistent within the allegedly changed climate; it remains an outlier (probably something like 5.1 sigma, i.e., 0.000017% probability) after an adjustment of μ , i.e., the same data point should be used to immediately reject the new hypothesis. A statistical alternative is increasing sigma, to incorporate the new data point, i.e., talk about "increased incidence of weather extremes", which is equally rejectable, since the rest of the data are nicely following the bell curve. Any Gaussian distribution function used will make that there are data points that should not have occurred for that distribution. A way out would be to say that the climate (μ and σ) has instantly changed, from one data point to the other. However, that goes against the common belief that climate changes are gradual. We are left with an inconsistent story that is representative for all empirical analysis and applies to all distribution functions used, and to the bell curve in particular.

The only distribution function that can be used is one that is equal to the distribution of data points that occurred, namely a sum of delta-dirac functions at the data points. This predicts the past data with complete accuracy, there where the bell-curve fails; the bell curve is in fact infinitely wrong in an infinite number of places, either non-zero where it should be zero and non-infinite where it should be infinite (Look again at Figure 1, the data are actually plotted differently than the model, each data line is a delta function).

This leads to strange results. If we have data points at 10.0 degrees, 10.2 degrees and 10.3 degrees, and make our correct triple-delta-function model, 10.1 degrees cannot occur? Well, it never did occur, so the best empirical model, by definition, is one that never predicts it. That is the definition of empirical forecasting. Only that what happened had a probability of happening. We seem to be sinking ever more in the empirical forecasting swamp. We mentally force the data into a nice and human-understandable probability distribution, to avoid any logical or computational analysis. (More about probabilities later). The bell curve masks our ignorance.

Continuing, empirical modelers often divide the data into two sets, a set to develop the model and a test set to validate the model. However, this does not help. The winning model will be that one that is validated best, so the one that describes the entire set best. What remains is that the better the past can be predicted, the better is the model, by definition. Up to and including abandoning the Gaussian distribution and alike functions (fat-tails, etc.), something that was never justified anyway, since it is not at all obvious that reality consists of multiple-coin-flip processes. Or, as Einstein said it "God does not throw dice". It is more likely that events in nature are the

result of dependent processes governed by rigorous physical laws, making the outcome deterministic and nothing like Gaussian at all. Reality is a delta-function 'distribution'; what happens happens with 100% certainty. These words were already uttered by Aristotle, "Only what will happen can happen" (Kenny, 2012). The fact that we do not know what exact delta function, does not make it Gaussian, even if often it looks Gaussian, which anyway it doesn't since reality is a sum of delta functions.

We'll get back to distribution functions later. For the moment the figure suffices to introduce the empirical reasoning used by most statistical analysts. Note that no physical reason is used to explain the outlier; the fact of 'lying out' is considered sufficient proof of climate change, as argued by the IPCC report. However, it is an empty statement, since any data point that moves into the moving averaging function mentioned above that is not equal to the value of the data point moving out of the window causes a change of average and thus μ (and possibly σ) and thus change of the climate, by its sheer definition. This explains why empirical predictions for the future always are mere extrapolations of trends. This happens in the climate and the stock markets alike. That is the fate of the empirical approach.

To give a financial example. Most insurance companies use bonus-malus systems for car insurance. If a person has had an accident this year, so the reasoning goes, the risk of that person apparently was underestimated, and that person has a higher probability of having an accident next year. They may even prove this reasoning by statistics and justify it mathematically by hindsight prediction (analyzing the 'roulette table'). However, in reality a person may well have a steady probability x to have an accident and any accident is ... well ... accidental. In that case it is just the probability estimate that has changed, because the model has been adjusted to account for the recent data. By looking only at the accident itself, without looking at its cause, the bonus-malus system is merely a scheme of delayed payment of the 'insured' person that had the accident, thereby eliminating any risk for the insurer. Modern insurance techniques use foresight prediction, basing the probabilities on factors such as the gender of the insured, which brings with it other moral complications such as the unjust generalization of groups of people, and moreover, the estimation of the risk parameters of the group members is normally based on ... yes, you guessed it ... empirical evidence.

Stochastic vs. deterministic

The Aristotelian quote of the previous section brings us to an important point. The difference between deterministic and stochastic processes. Statistics are a very nice but also dangerously complicated and a tricky instrument. Let us give you an example: We roll a 6-face die. Imagine

we roll it 1000 times and find the statistics. They are more-or-less as expected, with maybe a couple of times more a '3' compared to a '5', etc. But, let's assume that actually the die was completely regular and had the expected statistics with 1/6 probability for each number. That is, that we cannot exclude this possibility from our statistics of repeated throws, and we thus assume it because we like the symmetrical probabilities, i.e., we have a non-empirical model. Let's assume it is a probability function when throwing and that the dice is fully symmetric.

Now let's empirically analyze it. We roll it again, 3 times, and find '1', '6' and '4'. Have the probabilities of the die changed? Absolutely! If we only look at the last rolls, '6' and '4', we can say that the probabilities for these rolls have changed from (average)

1/6, 1/6, 1/6, 1/6, 1/6, 1/6

to (the last single rolls)

0, 0, 0, 0, 0, 1

And

0, 0, 0, 1, 0, 0

as proven by the statistics. Well, it happened, didn't it? Whatever the laws of nature underlying it, the final outcome was '4', so the best model is the one that predicts (retroactively) a '4'. Actually, both physically (understanding the laws of underlying physics) as well as empirically (just looking at the data). The real model, had we had knowledge of all of the underlying physics, would have predicted the above outcome. Continuing with 1/6 probabilities just shows our ignorance of the physical laws or initial conditions of the system. But, without knowing why the outcome was '4', the probabilities must have been 100% for a '4', because that is what happened. The system is deterministic to the bone. Now, for the future, the two approaches diverge: An empirical model will say that the next throw has high probability of '4'. An educated guess model will continue saying 1/6 for all numbers (because we philosophically like that because of symmetry and we keep it until we can reject this simple model), while nature itself will have a 100% probability for a certain number. What number we do not know yet, but the better we study it, the better we will know what number it is.

To see this better, take the extreme example of a deterministic but for a bystander seemingly random event as the Pharaoh shuffle of a deck of cards, a trick often used by magicians. For the ignorant audience it looks like

a random shuffle. The magician (steering the laws of distribution) makes it completely predictable by n times precisely interleaving the cards, which he learned to do repeatedly and reliably, i.e., predictably. Normal shuffles are less predictable by humans but equally deterministic by nature. The reasoning we use is: If I shuffle the deck in an equal way, the probability of card X being on top again is $1/52$. However, in reality, how we should think is: If I shuffle the initially identical deck of cards in an equal way under exactly equal conditions, the probability of card X being on top again is 100%. Something our magician does. The $1/52$ probability models our ignorance of the conditions and the way of shuffling.

The same is true for the weather. The weather events, like die-roll events and card shuffles, are completely deterministic. We are only blinded by our own ignorance into thinking there are some kind of probabilities involved. The fact that we cannot determine the physical laws of the system and calculate the outcomes, doesn't make it non-deterministic. As an example, the flows of air in the atmosphere and water in the oceans are turbulent and this is a stochastic process in our eyes. However, nature is deterministic – governed by the laws of thermodynamics – and only looks random to us. Thus, we are doomed to make wrong statements about the weather. It looks like a subtle difference, but the consequences are enormous, as will be shown.

The usual reasoning has an even greater flaw. It implicitly uses one essential axiom, something that is often forgotten. The distribution function used has to be of a probability type. That is, it should be integrable, and the integration should yield unity. That while, as we have just argued above, nature is deterministic. This limitation of functions is thus ungrounded; it is trying to force nature into a non-deterministic idea through probability functions, because we, mindless humans, only can work with it like that. Such functions thus get us into trouble by giving us misplaced confidence. We'll get back to function types later. Let's for the moment use the standard tactics of modeling the past data with docile probability functions and use them to make predictions for the future. Even then we get into trouble. That is because we use past data for parameter determination.

Let's go back to the situation of the definition of the climate as prediction of weather. Climate is the 30-point average of the weather. We could also use an example of the stock-market, where often a 50-day average is used to empirically see a trend and to invest (predict) on basis of this indicator, or we talk in technical terms of 'resistance levels' and 'support levels', in case of absence of any other, non-empirical, information. Just the data themselves. Now let's see what happens.

Take this simplified imaginary situation: The only information we have is this: There was never a flooding in Pakistan in 99 years of our

measurement set. Thus, the probability – our best-estimator for the risk for the next year – of flooding in Pakistan was 0%. Mind you, we use statistics only, and do not include any knowledge of physics whatsoever in our prediction model.

Then Pakistan gets hit by flooding. What is now our best-estimator of probability of flooding in the next year? 1%? Namely, once every 100 years? Depends on how the climate is defined! Our reasoning may go something like this:

- The probability of flooding was 0% and the occurrence of flooding was 0% until last year, as proven by facts. Our best model for the past was 0% probability.
- The probability of flooding (and occurrence of flooding) last year was 100%, as proven by facts. Without understanding why, our best model of last year was 100% probability.
- The probability of flooding for the coming years is 3.3%. Namely, it happened once in the thirty-year climate window. The climate is now defined as "Flooding? Once in thirty years, i.e., 3.3%." Any other conclusion is a wrong interpretation either of the statistical data or of the definition of climate.

Note that the probability of a flood (in the coming years), depends on our quite arbitrary definition of climate. $p = 1/n$, with n the number of years used for the definition of the climate. Part of the events are 'swept up' into climate, and part of it in the variability within the climate, i.e., weather, and the demarcation line between them chosen quite arbitrarily. If we don't understand the physics behind the system, and use only the data, we naturally wind up with this conclusion. The probability of a flood tomorrow depends on the (arbitrary) definition of climate. That is absurd, but is the only possible outcome of the empirical approach.

To show how weird this is, imagine we were to do this with the roll of a die, if the probabilities of the die were purely empirically determined by the last thirty rolls. This seems ridiculous to us, since we do have a physical explanation for the probabilities of this system. A die has octahedral symmetry (called O_h), and we know it has 6 symmetrically-equivalent minimal-gravitational-energy positions (plus 8 maximums and 12 saddle points). That makes us ignore the 30-roll-statistics and stick to our non-empirical probabilities, since we deem our derived physical/mathematical laws of higher order than the actual data. (But this also seems wrong, since it cannot even retrodict the past. What model is that?!)

For the stock-market, or the climate, or any poorly-understood system, for lack of something better, we use an empirical approach of data only. To be able to say if the market (bull or bear market) or climate has changed we look if the average in the past data has changed. From a human

point of view the system is described as best-estimators of market/weather/climate parameters (price, temperature, precipitation, etc.) and they are the averages of those parameters in our averaging window, any event that has a value for a particular parameter that is not exactly equal to the current average value will change the average and thus our best-estimator for that parameter and thus our estimation of the market or climate. "Did the climate change?" is thus converted into nothing more than "Has the average value changed?" This is a trivial result of our definitions. The only discussion we can now have is about the direction and magnitude of changes and eventually about the permanency of these changes, where the latter cannot be decided on statistics alone.

Meanwhile, empirical agencies find correlations in and between data and (re)invent an ad hoc explanation for them and use the hindsight predictiveness of the model to prove the correctness of the model. Ironically, applying the empirical method to these agencies themselves – the failed past predictions as input 'data' – will make us completely ignore them altogether. It is just a matter of empirical analysis; the probability they are right this time is estimated to be zero, since they have never been right before in the past.

Coming back to the climate and weather events, if we abandon the empirical track, and study more, namely the underlying physical processes, we can determine what it is exactly that has changed in the climate, why the event occurred. But that the climate has changed is irrefutable. Any statement about the climate on basis of a past weather event is therefore completely void of information. Yet, it appears in literature over and over again (Stott, Peterson, and Herring, 2012).

The more we study and understand the physical laws governing the weather, the more we stop talking about probabilities and start talking about certainties: "It flooded because of X, Y and Z. These things were absent in the years before, and last year they were there. So, the system – the climate – has changed, namely in the aspects X, Y and Z". If we had the mother-of-all-computers, combined with an infinite knowledge of physics and facts, including all butterflies in Buenos Aires or whatever tiny event on our planet and the rest of the universe, we would also see the weather as something deterministic. There is no reason to believe that quantum mechanical fundamental Heisenberg uncertainties influence the large-scale physics and make weather and climate non-deterministic (and even there the final word has not been said yet). In fact, Chaos Theory – very relevant for the weather – is the embodiment of the idea that nature is deterministic but too complex for humans to understand, it is pseudo-random. We cannot see the difference between a stochastic and a deterministic system, and to nearly all effects they are the same, but they are fundamentally different. Nature is deterministic and our probabilistic models have no base. These ideas can be transposed to

other systems. For example, the similarity of weather and stock markets is plain enough, as both are chaotic fractal systems based on intricate feedback processes.

Distribution functions:

We have already briefly mentioned the aspect of the type of function used for statistics in the previous section. There is no logical base for requiring probabilistic functions at all, and even less so for using the Gauss-like curves for past and future predictions of all systems. In this part we criticize specifically the use of the bell curve. Others have gone before us. As, for instance Fendler and Muzaffar write, "Our purpose is to criticize [...] the belief that a normal-curve distribution is a representation of real things in nature" (Fendler and Muzaffar, 2008). For some systems it might work. For others blatantly not. To give you an example where it doesn't fit: people's performances. Until recently thought to be Gaussian, the real distribution of the performance of people has a clearly non-Gaussian long tail (O'Boyle Jr. and Aguinis, 2012). To give an example where it does fit reasonably well, take again the die rolls. This is a nice docile system. The number of possibilities is finite (exactly 6), and even if it is a biased die, we can predict the die roll to some degree, and cannot expect outliers.

It boils basically down to this idea. For nice docile systems like die rolls, statistics can be used. We can make a prediction on the average – the expectation value – of future die rolls. Actually, if we are knowledgeable, we can 'predict' the die roll with certainty, just as we can predict card shuffles, especially of the Pharaoh type. For this type of systems, our level of knowledge is represented in the level of our statistical approach. This then ranges from completely ignorant (empirical; functions and parameters are derived from data themselves) to mildly ignorant (functions and parameters are found on basis of reason, like the die), to fully sapient, where there is no 'distribution' at all, but prediction entails just working out the calculation for the future, like in Pharaoh shuffles and billiard ball collisions.

Note, however, that we have to be at least sapient enough to know that we can use the probability-functions-statistics approach in the first place. We have to know that the system has an expectation value to start with. This is not the case for all systems. While the next die roll has an 'average', an expectation value, other systems do not have this quality. For example, it is not possible to say anything about the expectation value of the next earthquake in San Francisco. That is because these do not follow probability-like statistics. Remember that a probability function was not required for deterministic systems and that nature is a fully deterministic system and thus not restricted by our simplistic mathematics. Earthquakes, and many natural systems, do not follow Gauss or Gauss-like distributions, but other, much

more troublesome functions instead. This changes the game completely, as we will exemplify using the climate again.

In a completely-ignorant way of analyzing the system, the idea is that some weather parameter (like temperature or rainfall) is assumed to be normally-distributed with a Gaussian probability function (Eq. (1)), see Figure 1. If an event then lies more than 2σ or 3σ away from the median μ , it is considered abnormal and thus a sign of climate change. Although any average-changing event is climate change – 'outliers' as well as 'inliers', as discussed above – this particular reasoning is false.

The reasoning normally followed is this:

1. Assume the events are normally distributed with a Gaussian probability distribution function (Eq. (1)).
2. Find the parameters μ and σ best representing the real data.
3. Determine by a Student's t-test if a new event lies within 2σ or not and if it is consistent with the null-hypothesis "parameters unaltered" or is an extreme event. Fig. 1, has a 5.4σ for outlier of 2004. If the event is extreme, the distribution (climate) must have changed.

The reasoning goes wrong with the basic assumption, namely of the distribution being Gaussian. Upon arrival of a new datapoint that is falling outside the expected range of the normal distribution, it is then concluded that the parameter μ has changed (climate change), but instead the conclusion should be that the data are not distributed normally. If a datapoint does not fit well with distribution function F and parameters P , it might mean that the parameters P are different or changed – as we are wont to believe – but it might as well be that the choice of function F is wrong. In fact, it might not even be a probability-like function at all. Especially so, since we have just seen that many systems are not at all about probabilities.

The problem lies in the fact that it is as an axiom assumed that the weather follows a probability distribution, such as Eq. 1, that integrates to unity, while, as reasoned above, the system does not have probabilities. What remains is only a frequency-of-occurrence tallying statistics, so called 'binning'. These can have strange properties indeed, as we shall explain in more detail.

There are basically two types of functions $f(x)$ to describe the tail behavior (x far away from median) of statistics in nature, namely exponential $n^{-y(x)}$ (as the tails of the Gaussian bell curve, with $n = e$ and $y = x^2$), and power-law x^{-n} , also called Pareto functions, Zipf law, etc. Members of the first family of functions – also called 'non-scalable' – are very 'docile' and will not cause the scientist much problems, and thus the primary reason for their use, "fooled by the reductionist need for simple functions" (Taleb, 2011). There are also systems that follow the second family of functions, also called 'scalable', defined as the fractal property that any multiplication

of x to ax introduces a constant factor in the function value, independent of x (Taleb, 2011),

$$\frac{f(ax)}{f(x)} = \frac{f(ax')}{f(x')} \quad (3)$$

For instance, an earthquake of Richter (logarithmic) magnitude x has a times more occurrence than an earthquake of magnitude $x + 1$, independent of x . It is easy to show that the Gaussian is not scalable. Which functions are? The above equation can be solved to find the set of functions that are scalable. In the first step, noting that the fraction is constant, we take the logarithm on the left side, divide it by $\ln(a)$ and letting this factor go to zero,

$$\lim_{\ln(a) \rightarrow 0} \frac{\ln[f(ax)] - \ln[f(x)]}{\ln(a)} = C. \quad (4)$$

The left side here is the definition of a derivative of the function on a log-log scale ($\ln(ax) - \ln(x) = \ln(a)$). Therefore, the function in a log-log plot is a line, with constant derivative,

$$\frac{d \ln[f(x)]}{d \ln(x)} = C,$$

$$\frac{d \ln[f(x)]/dx}{d \ln(x)/dx} = C,$$

$$\frac{d \ln[f(x)]}{d(f(x))} \cdot \frac{d(f(x))}{dx} = C,$$

$$\frac{1}{f(x)} \cdot \frac{d(f(x))}{dx} = C,$$

$$\frac{d(f(x))}{dx} = C \frac{f(x)}{x}. \quad (5)$$

This is a differential equation which indeed has as unique solutions the power-law functions mentioned before,

$$f(x) = f_1 x^{-\alpha}, \quad (6)$$

with f_1 a constant being the function value at $x = 1$ and α a constant ($\alpha = -C$) denoting the power (for instance, earthquakes follow power $\alpha = 3.04$ (Newman, 2005)) – the lower the number, the more the system is pestered by outliers (Taleb, 2011). (See Figure 2 for an example with $\alpha = 0.5$). 'Scalable', 'fractal' and 'power-law' are all one and the same thing. See the work of Newman for an excellent summary of these power-law functions (Newman, 2005). Power-law functions can behave like Gaussian functions, especially for small numbers of samples, but they lack the parameter σ and drawing conclusions on basis of this parameter can thus not be done.

Actually, these functions are horrendously difficult to work with. It is as good as impossible to draw conclusions with these functions, as anybody working with them knows. And even if it is not a power-law, that does not mean that the distribution is as simple as normal Gaussian (or whatever other function we assumed). Some distributions are (stretched) exponential $\exp[-(x/x_0)^\beta]$, equally difficult to work with and therefore equally avoided by scientists trying to model the world and predict the future. In the worst case, we can have non-parametric statistics that are not described by any well-known function with parameters. Even if we cannot come up with something better, and that most of the times it may look correct, our choice of Gaussian is not necessarily correct. Any outliers instantly prove us wrong, the moment they appear.

The function of probabilities, for instance for the lottery, is by definition integrable (and integrated to unity; a limited-range number has to come up and somebody has to win). Not so for the frequency-of-occurrence functions of many natural phenomena. A power-law function (Eq. 6) cannot be integrated over all possible values of x (zero to infinity), since it results in infinity, for any value of α , either on the zero-side or on the tail side of x . (Often, arbitrary lower or higher limits are set to x , for instance the strongest earthquake ever measured, until now, to force it to be usable, integrable, because that is so much needed). It can thus also not be scaled to result in unity when integrated. 'Probabilities' are not defined; only real-events statistics remain. As, such, the average expected magnitude of a future event,

$$\langle x \rangle = \frac{\int_0^\infty xp(x)dx}{\int_0^\infty p(x)dx}, \quad (7)$$

cannot be calculated on basis of empirically-found parameters (f_1 and α). In other words, the average magnitude is not defined for such phenomena (ex. earthquakes; One cannot make a statement about the predicted average of the next earthquake). There are some once-in-a-while-occurring high-magnitude outliers – Black Swans – that mess-up the analysis. Since, on basis of its fractality, the weather is also expected to be a scalable (power-law) system, one cannot make an estimation for the average temperature in, say, Amsterdam. One can only make a statement about the past average temperature there. The past has an average, the future not (yet). Outliers are natural and they mess up the calculations.

Thus, the bell-curve is not the best choice of function. As an example, flow of matter (air and water), a fundamental process for the weather and the climate, shows turbulent behavior and thus results in fractality and scalability, i.e., power-law. Likewise, many financial markets are composed of an intricate network of coupled communicating nodes,

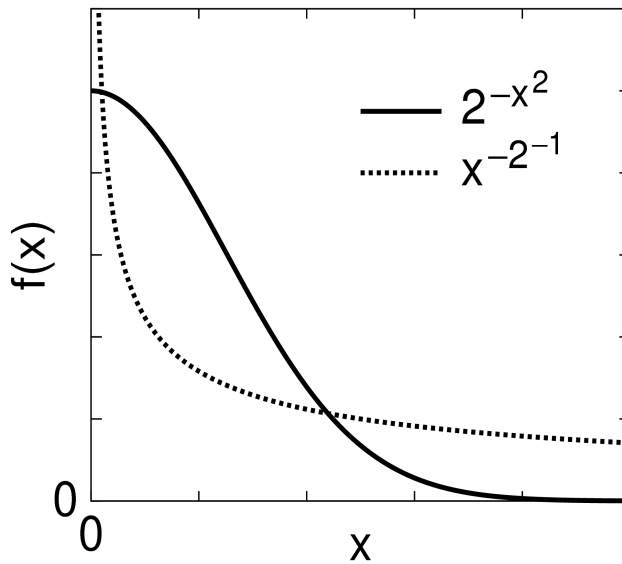


Fig. 2: Two types of functions (half shown, only positive x), non-scalable (like Gaussian, or the similar 2^{-x^2} shown with solid line), and scalable, power-law (like the shown $1/x$, dashed line). The former is docile and integratable, the latter not (the shown example because of the long tail, not going fast enough to zero; for other exponents, problems can occur at $x = 0$)

causing chaotic behavior. These are fractal-like systems, where zooming-in reveals basically the same or similar patterns. They are therefore probably not well described by (docile) probability functions like the Gaussian, but should rather be modeled with Mandelbrotian functions (Taleb, 2011) like power-laws, even if that makes our work difficult, if not impossible. Yet, if we, scientist, model the system with bell-curves, and find that data are not consistent, we prefer to throw away the data – call them 'outliers' – rather than throwing away the model, or use them at our convenience to make statements about the changeability of the system. That while in most cases, we have only the data and no underlying physical knowledge of the system to work with.

Scientists, in an attempt to describe nature, use stochastic models which necessitate the use of probability-like docile functions, i.e., Gaussian bell curves. If, in this way, we manage to successfully describe 99% of the real situations we are happy. The remaining 1% is just 'outliers' and does not discourage us; The 99% success rate validates the stochastic model approach.

Extrapolation

The next step conventionally done is a time-series analysis. After a statistical analysis of finding μ and σ , one can do a linear regression on the data and try to find a tendency, trying to find out how the data depend on time t , i.e., finding a and b in the equation $\mu(t) = at + b$, rather than only

analyzing their statistical distribution. It is then very tempting to try to do an extrapolation and make a prediction for the future. See for example Figure 3. In this figure, a clear tendency of temperature dropping is visible (it was actually introduced in the simulation, 0.2 degrees per decade). So, it is very straightforward to draw a conclusion that next year it will probably be even colder. However, general wisdom is that extrapolation is not allowed on basis of data alone. No linear extrapolation, unless we understand the physics behind the system and have explained that it should be linear. Never on basis of data alone. In the figure, what conclusion can we draw? Will it be cooler next year? No, the only conclusion we can draw that it has cooled down in the previous years. This seems counterintuitive, but is the only reasonable conclusion.

In his book "Thinking, fast and slow", Kahneman explains the extrapolation fallacy in detail (Kahneman, 2013). More specifically in the chapter called "Regression to the average": Imagine a golf tournament. A player's score is determined in part by his skills, but also to great extent by his luck. Thus, if after one day a certain player is on the first place of the leaderboard, it is because he is skilled and he had exceptional luck. While his skills will not have dwindled much overnight, his luck is expected to be average next day, and thus we expect him to drop on the leaderboard. This is counterintuitive and most people will expect a person leading a tournament to continue to play well and continue leading, in fact extend his lead. That is because people use an empirical analysis, extrapolating past performance for a prediction of the future. However, theory tells us and reality shows us that if tomorrow's performance is less than 100% correlated to today's performance (not uniquely determined by skill), regression to the average is to be expected. The less the correlation, the more the system will regress to the average. Note that a sports event where luck is banned is as exciting as watching a dry tumbler in action. Generally speaking, we watch sports exactly because it is not fully predictable.

This same type of reasoning may also be applied to the climate. In 1998 we had an exceptionally warm year. This was due to climate change (equivalent to 'skills' of golf player), but also due to random variations ('luck' of golf player). The conclusion of the people using an empirical approach only was that it would warm up even more. Yet, regression to the average told us that after an exceptional warm year it was bound to cool down. As indeed happened. Gosselin compares the real data of 2014 with the predictions for 2014 published by Smith in Science in 2007 (Smith et al., 2007), apparently "a staggering failure" (Gosselin, 2014); we are well outside any error margins of their empirical models, we have regressed to the mean of linear warming up that started in the 17th century.

To show where simple data analysis and extrapolation may go wrong, consider the case of a melting of an ice sheet, say, for instance, the Greenland ice cap. The same Figure 3 shows an imaginary scenario with a saw-tooth function of the sheet thickness. This is not so far-fetched, as Greenland is one of the biggest deserts on our planet ("In fact there is less rainfall in Northeast Greenland than in the Sahara, and thus the expression 'the Arctic desert' has arisen" as stated by the Greenland tourism agency (greenland.com, 2015)). Precipitation there is also very irregular; It may not snow for 50 years, and then suddenly snow a lot (Johannessen et al. reported indeed an unexpected growth of the ice-sheet in Greenland in 2005 (Johannessen et al., 2005)). Climate scientists, monitoring the ice sheet thickness, will always measure a declining ice sheet, yet, over time, nothing changes, as can easily be verified in the figure. The thing is, if we do not understand the physics behind the system, but only look at the pure data, we are doomed to make erroneous extrapolations.

Correlations

A third powerful tool of empirical statistics is correlations, often used to prove theories. Examples are plenty. Famous examples can be found in medicine, economy and once again the climate.

In medicine, it was shown beyond reasonable doubt that a correlation

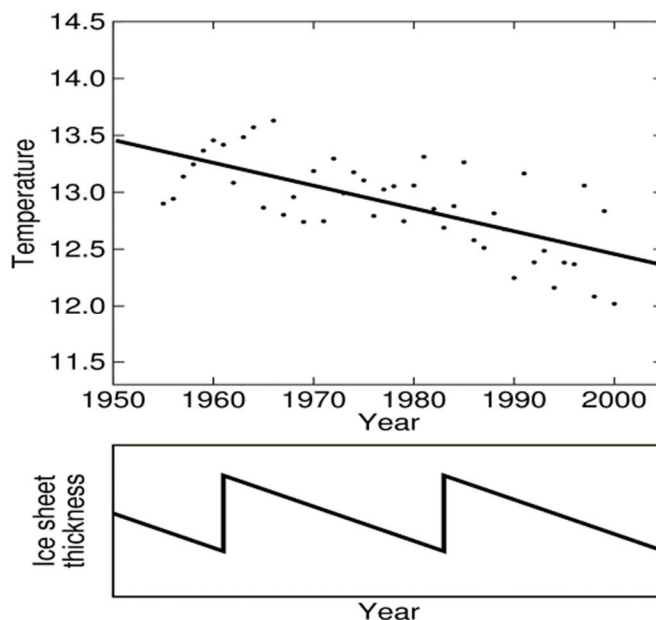


Fig. 3: Extrapolation of data. Top: A set of data of random (homogeneous, 1 degree spread) on a declining temperature (2 degrees per decade). Are we allowed to extrapolate? Bottom: A simulation of ice sheet 'melting'. At any point the ice disappears, however, on the long scale nothing changes

exists between high cholesterol levels and cardiovascular problems. Pharmaceutical industry used this correlation to sell its idea that their cholesterol-lowering medicine (statins) would be beneficial for the public. No positive effects ever materialized. Thus, an alternative explanation might exist, namely that high cholesterol levels are caused by a risk of cardiovascular disease, the body fighting the risk by increasing cholesterol levels. An hypothesis that was never studied, since obviously no money can be made on it.

In climate research, a high correlation was found between atmospheric carbon-dioxide concentration and global temperature (Gore, 2006). However, this correlation was only seen in the past, while recent huge amounts of CO₂ injected into the atmosphere apparently were accompanied by a lowering of global temperature (since 1998). In the economy a correlation was shown to exist between government debt and stagnation of the economy (Reinhart and Rogoff, 2010). The banking sector – for fear of losing its investments – used this to convince governments to implement austerity measures, and five years later, no positive effects of these measures materialized. Quite to the contrary. Economy slumped everywhere.

Pro-oil research established a correlation between oil prices and misery (and the more common idea of wealth and energy consumption per capita) (Skov, 1999).

The main problem with correlations without a sound scientific explanation is that the empirical data do not show the arrow of cause and effect. Correlation is not causation. Which one is causing which? They can be even completely coincidental, or have a common underlying reason. Assuming causation is a logical fallacy called *cum ergo propter hoc* (Latin for "with this, therefore because of this").

There are basically six types of resulting correlations between A and B, see Figure 4:

1. Coincidental: A and B independent of each other, but by chance seem to be correlated
2. Common base: X causing A and B. X is normally called a 'confounder'
3. Transitive (very common): A causing X causing B, A not causing B
4. Combinative: A plus X causing B, A or X independently not causing B. (ex: Water plus cold gives ice)
5. Unidirectional: A causing B or B causing A
6. Bidirectional (feedback): A causing B and B causing A

The last one is the most complex and has the most chaotic behavior. Note that the causalities in the two directions in that case do not necessarily have to be of equal strength (if A and B have the same unit).

As an example, upon scrutiny, it was shown that CO₂ is lagging behind temperature variations by some thousands of years. Indermühle and coworkers (Indermühle et al., 2000) made a statistical analysis and found a value of 900 years for the delay and note that "This value is roughly in agreement with findings by Fischer et al., who reported a time lag of CO₂ to the Vostok temperature of (600 ± 400) yr during early deglacial changes in the last 3 transitions glacial-interglacial" (Fischer et al., 1999). That makes the correlation of type 5 possible, but only if CO₂: coincidental is the result of temperature and not the other way around. In view of this, CO₂ capping is fruitless, but can devastate human civilization by reducing the benefits (!) of CO₂, high level of industry and wealth (Earlywarn.blogspot, 2011), high level of agriculture by CO₂ fertilization (Hartwell Allen Jr., Baker, and Boote, 1996), etc.

Likewise, in the economy, it was argued by Krugman that actually the high government debt was the result of a stagnating economy (Krugman, 2013a, Krugman, 2013b) and not the other way around, as earlier suggested by Reinhart and Rogoff (Reinhart and Rogoff, 2010). With this idea, austerity can wreck the economy even further, down-spiraling into economic disaster, as we now see taking place in front of our eyes in many countries around Europe. In our country, Portugal, the Austerity measures made the government debt rise from 94% of GDP in 2011 to a staggering 129% of GDP in 2014, even though the European Union started including prostitution and illegal drug sales in the GDP, everything to get the GDP up (and the debt-to-GDP ratio down). It seems obvious, while there is a correlation, in fact stagnation causes a public debt (ratio) and not the other way around.

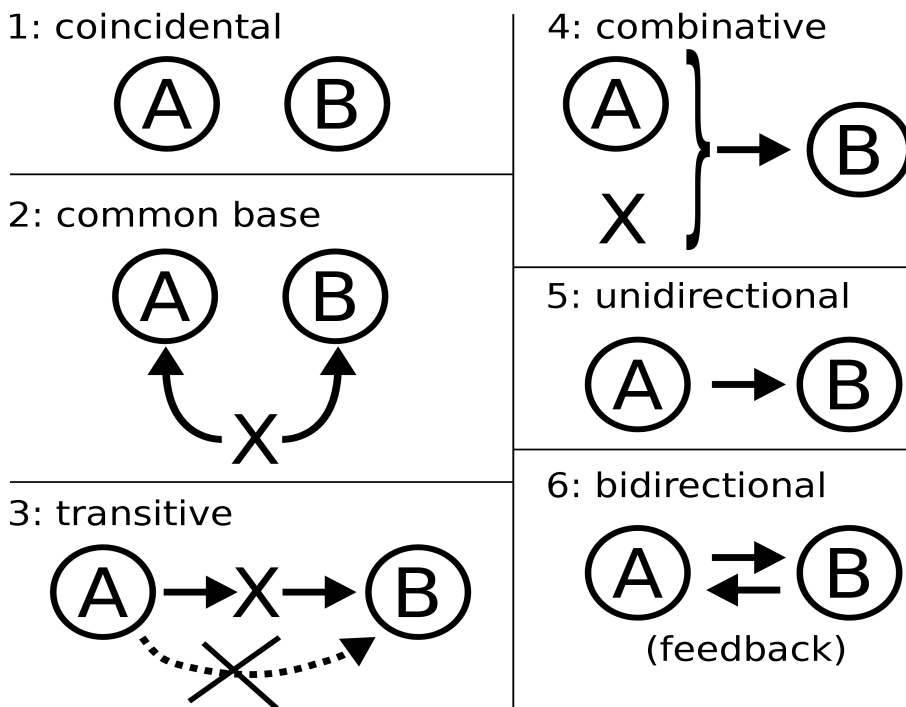


Fig. 4: Six different ways of correlations 1: coincidental, 2: common base, 3: transitive, 4: combinative, 5: unidirectional, 6: bidirectional (feedback)

For cholesterol, it might be so that both high cholesterol and cardiovascular diseases find a common basis in the DNA of the person. Or it might be so that high cholesterol is the only way the body can lower the high risk of heart attacks. In the former scenario, cholesterol-lowering statins will have no effect whatsoever, as seems to be the case. In the latter it will even be counterproductive, since we disarm the body in its fight against illnesses. Cholesterol might actually be good for you! (Schatz et al., 2001); (Kendrick, 2008) And, actually, obesity might not even be a problem for a population, as statistical data indicate (Gibbs, 2005). The same way in the 1950’s doctors balanced the fight of the body with infections in favor of the bacteria – based on irrefutable correlations between low-iron levels and infections – by giving iron supplements to sick patients, thereby factually killing them. Note that bacteria need iron much more than the body does (Schaible and Kaufmann, 2004); our body secretes iron in order to kill the bacteria. The myth of the benefits of iron was further amplified by the comical character Popeye eating spinach, thus combining two myths in one, since apart from iron having no benefit to an infected person, also spinach does not contain so much iron as previously claimed. (To add to the confusion, this myth might be a myth. Popeye was based on vitamin A in spinach (Sutton, 2010)).

These examples show where failed measures are based on failed research, neglecting the scientific method – for instance the requirement to show that the model is the only one that can explain data (Stallinga and Khmelinskii, 2013). To further show you how this works, imagine politicians listening to scientists that analyze health of people. We know that men smoke more than women and that smoking is bad for our health. The obvious conclusions of scientists will be that being a man is bad for a person's health and politicians implement mass-penectomy measures on basis of these scientific facts. All because they do not understand correlations of type 3 above (A causing X causing B, A = 'man', X = 'smoking', B = 'cancer', where each step actually can, in turn, be transitive itself) and only look empirically at correlations that tell us nothing if the scientific method was not applied and alternative explanations were not included. The literature is littered with examples of these non-scientific statistical correlations, more so since they are the easiest to publish (since models that are consistent with data are indeed difficult to reject by the referees). The conclusion is that if we use an empirical approach, analyzing the data without a good theory to explain them, correlations can and will fool the scientist. This non-scientific empirical approach will inevitably have deleterious social and political consequences if used in decision making. The correlation between CO₂ and global temperature used by politicians to justify flagellation of the world by misplaced measures maybe the best example of how the entire society can be suffering from non-scientific research.

Conclusion

In conclusion, we argued that if we use a purely empirical data-only analysis, we can do this only of the past data and any predictions for the future are meaningless. To make predictions for the future on basis of past data, we have to know what the probability function is – if the system is governed by a probability function at all. That is, we have to at least know something, whereas deriving the distribution function and its parameters on basis of the data alone is not useful and even misleading. In case the distribution function is not probabilistic, like the power-law observed for many phenomena, no prediction of the future average – no expectation value – can be calculated.

Likewise, if we want to make extrapolations based on trends, we are only allowed to do this if we have a good basis for our extrapolation function. In other words, whatever we do, purely empirical forecasting is meaningless. An example given was the climate. After an extreme event ('outlier') it is often said, "This event was extremely improbable and is (thus) a sign of climate change". The word 'thus' is imagined by the public or inserted by journalists somewhere in the process, as no self-respecting

scientist will utter it. Weather and climate do not follow probability distributions, but only reality frequency-of-occurrence distributions and thus any probabilistic statistical analysis is meaningless. The second part – 'climate change' – is an empty statement, since anything that is not equal to the past average is a sign of climate change by the sheer definition of the climate (as the 30-year weather average). What remains is, "The event had small or zero occurrence before (but occurred now, it is an outlier)". On top of that, the outlier used in the argumentation often remains an outlier in the Bayesian-adjusted model and the new model should be rejected as fast as the old one.

These ideas presented here are easily applicable to any empirical-forecasting system, such as any financial markets, or sports events, or disease epidemics. Finally, it was argued that blind analysis of correlations does not make sense. All in all, we must conclude that the empirical approach is a simpleton's tool, and a scientist's nightmare. Any analysis without an understanding of the underlying processes is full of pitfalls. The approach of the scientist represents the level of understanding of the system, ranging from completely ignorant (fully empirical; everything based on past data) to completely sapient (prediction is just working out the equations of the physical laws). In the former case, we will unavoidably get into trouble sooner or later with our predictions, as history has shown us over and over again.

Yet, as a positive note to end, we should not be too discouraged and abandon the empirical approach altogether. Inductivism is part of science and the community is benefited by a wide range of creative and imaginative ideas and theories. The fancier the better. The only thing is that we should not be too surprised if our ideas fully miss the mark. An empirical approach is bound to make us look like fools most of the times. Eventual winning theories will have passed heavy falsification. The others we will forget altogether.

Disclosure of interests

The authors declare not to have any conflict of interests. This research was paid by no grant. It received no funding whatsoever, apart from our salaries at the university where we work. Nor are we members of any climate committees (political or other) or are we linked to companies or NGOs, financially or otherwise. We are not members of any political party or movement. This is an independent work that does not necessarily represent the opinion of our university or of our government.

References:

1. Abounoori, E., M. Shahrazi, and S. Raseki (2012). “An investigation of Forex market efficiency based on detrended fluctuation analysis: A case study for Iran”. In: *Physica A* 391, p. 3170. doi: 10.1016/j.physa.2011.12.045.
2. Angiulli, F., V. Fionda, and S. E. Rombo (2007). “Protein Data Condensation for Effective Quaternary Structure Classification”. In: *Proc. 8th Int. Conf. Intell. Data Eng. and Autom. Learning (IDEAL 2007)*, 810 and LNCS 4881. doi: 10.1007/978-3-540-77226-2_81.
3. Arguez, A. and R. S. Vose (2011). “The Definition of the Standard WMO Climate Normal”. In: *Bull. Amer. Meteor. Soc.* 92, p. 699. doi: 10.1175/2010BAMS2955.1.
4. Chalmers, A. F. (1982). *What is this thing called science?* 2nd ed. Open University Press. ISBN: 0-335-10107-0.
5. Chen, S.-H. and C. H. Yeh (2002). “On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis”. In: *J. Econ. Behavior Org.* 49, p. 217.
6. Cordier, S., L. Pareschi, and G. Toscani (2005). “On a Kinetic Model for a simple Market Economy”. In: *J. Stat. Phys.* 120, p. 253. doi: 10.1007/s10955-005-5456-0.
7. Earlywarn.blogspot (2011). url: <http://earlywarn.blogspot.com/2011/02/energy-and-wealth-arent-that.html> (visited on 02/10/2011).
8. Fendler, L. and I. Muzaffar (2008). “The history of the bell curve: sorting and the idea of normal”. In: *Educ. Theory* 58, p. 63.
9. Fischer, H. et al. (1999). “Ice core records of atmospheric CO₂ around the last three glacial terminations”. In: *Science* 283, p. 1712. doi: 10.1126/science.283.5408.1712.
10. Gibbs, W. W. (2005). “Obesity: An Overblown Epidemic?” In: *Sci. Am.* (June), p. 48.
11. Gore, Al and Davis Guggenheim (director) (2006). *An Inconvenient Truth*. Motion picture.
12. Gosselin, P. (2014). *Laughing Stock Met Office...2007 "Peer-Reviewed" Global Temperature Forecast A Staggering Failure*. url: <http://notrickszone.com/2014/06/24/laughing-stock-met-office-2007-peer-reviewed-global-temperature-forecast-a-staggering-failure/> (visited on 06/24/2014).
13. greenland.com (2015). *Wind, Snow and Rain*. url: <http://www.greenland.com/en/about-greenland/nature-climate/the-weather-in-greenland/wind-snow-and-rain/> (visited on 12/08/2015).

14. Guemas, V. et al. (2013). “Retrospective prediction of the global warming slowdown in the past decade”. In: *Nature Climate Change* 3, p. 649.
15. Hartwell Allen Jr., L., J. T. Baker, and K. J. Boote (1996). “Global climate change and agricultural production. Direct and indirect effects of changing hydrological, pedological and plant physiological processes”. In: *Food, Agriculture Organization of the United Nations, and John Wiley & Sons. Chap. The CO2 fertilization effect: higher carbohydrate production and retention as biomass and seed yield.* isbn: 92-5-103987-9. http :
//www.fao.org/docrep/w5183e/w5183e06.htm.
16. Holbrook, S. R., S. M. Muskal, and S.-H. Kim (1993). “Artificial intelligence and molecular biology”. In: *American Association for Artificial Intelligence Menlo Park, CA, USA. Chap. Predicting Protein Structural Features With Artificial Neural Networks*, pp. 161–194. isbn: 0-262-58115-9. doi: 10.1.1.101.430.
17. Indermühle, A. et al. (2000). “Atmospheric CO2 concentration from 60 to 20 kyr BP from the Taylor Dome ice core and Antarctica”. In: *Geophys. Res. Lett.* 27, pp. 735–738.
18. IPCC (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.*
19. Johannessen, O. M. et al. (2005). “Recent Ice-Sheet Growth in the Interior of Greenland”. In: *Science* 310, p. 1013. doi: 10.1126/science.1115356.
20. Kahneman, Daniel (2013). *Thinking, fast and slow.* Farrar, Straus and Giroux. isbn: 978-0374533557. Karlis, D. and I. Ntzoufras (1998). “Statistical modelling for soccer games: The Greek League”. In: *Unkown.*
— (2003). “Analysis of sports data by using bivariate Poisson models”. In: *The Statistician* 52, p. 381.
21. Kendrick, M. (2008). *The Great Cholesterol Con.* John Blake Publishing Ltd. Isbn: 978-1844546107.
22. Kenny, Anthony (2012). *A New History of Western Philosophy.* Oxford University Press. Isbn: 978-0199656493.
23. Krugman, P. (2013a). “Debt and Growth: The State of the Debate”. In: *The New York Times* 31 May. url:
24. <http://www.nytimes.com>.
— (2013b). “The Excel Depression”. In: *The New York Times* 18 April. url: <http://www.nytimes.com>.

25. Lovejoy, S. (2014). “Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming”. In: *Clim. Dyn.* 42, pp. 2339–2351. doi: 10.1007/s00382-014-2128-2.
26. Lynch, P. (2008). “The origins of computer weather prediction and climate modeling”. In: *J. Comp. Phys.* 227, p. 3431. doi: 10.1016/j.jcp.2007.02.034.
27. Newman, M. E. J. (2005). In: *Contemp. Phys.* 46, p. 323.
28. O’Boyle Jr., E. and H. Aguinis (2012). “The best and the Rest. Revisiting the Norm and Normality of Individual Performance”. In: *Personnel Psychology* 65, p. 79. doi: 10.1111/j.1744-6570.2011.01239.x.
29. Reinhart, C. M. and K. S. Rogoff (2010). “Growth in a Time of Debt”. In: *Am. Econ. Rev. Papers Proc.* 100, p. 573. doi: 10.1257/aer.100.2.573.
30. Schaible, Ulrich E. and Stefan H. E. Kaufmann (2004). “Iron and microbial infection”. In: *Nat. Rev. Microbiol.* 2, pp. 946–953. doi: 10.1038/nrmicro1046.
31. Schatz, I. J. et al. (2001). “Cholesterol and all-cause mortality in elderly people from the Honolulu Heart Program: a cohort study”. In: *The Lancet* 358, p. 351. doi: 10.1016/S0140-6736(01)05553-2.
32. Skov, A. M. (1999). “National health, wealth, and energy use”. In: *Soc. Petroleum Eng. and JPT* May, p. 48.
33. Smith, D. M. et al. (2007). “Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model”. In: *Science* 317, pp. 796–799. doi: 10.1126/science.1139540.
34. Stallinga, P. and I. Khmelinskii (2013). “THE STATE OF SCIENCE IN 2013”. In: *European Scientific Journal* 4, p. 385.
— (2014). “Consensus in science”. In: *Monte Carlo Meth. Appl.* 21. doi: 10.1515/mcma-2014-0008.
35. Stott, P. A., T. C. Peterson, and S. Herring (2012). “EXPLAINING EXTREME EVENTS OF 2011 FROM A CLIMATE PERSPECTIVE”. In: *Am. Meteo. Soc. July*. Ed. by T. C. Peterson, P. A. Stott, and S. Herring, p. 1041.
36. Sutton, M. (2010). “Spinach and Iron and Popeye”. In: *Internet J. Criminol.* P. 1.
37. Taleb, N. N. (2011). *The Black Swan: The Impact of the Highly Improbable*. Allen Lane. Isbn: 978-1846145650.
38. Timmermann, A. and C. W. J. Granger (2004). “Efficient market hypothesis and forecasting”. In: *Int. J. Forecasting* 20, p. 15. doi: 10.1016/S0169-2070(03)00012-8.

39. Watts, Anthony (2013). Watts Up With That? url: <http://wattsupwiththat.com/2011/07/16/i-need-your-help-for-a-short-research-project/> (visited on 07/15/2013).