

Recurrence Statistic Indicator v4.2

The main idea:

Forex price signal produces trends only rarely, when there is a great imbalance between supply and demand. Most of the time the price retraces itself in a relatively narrow area of price.

This observation was put in the mathematical language by Eurussd as following:

Definition: A price $X_T(t_0)$ is $h(T)$ recurrent whenever $X_T(t_0)$ is between the high and low of the bar in the timeframe T , then at least one of the previous or next h bars passes through $X_T(t_0)$.

Proposition 1: If $X_T(t)$ at any time t relative to Timeframe T , then almost surely, there exist positive integer h and k such that every price belonging to the set $[X_T(t)-k, X_T(t)+k]$ is $h(T)$ recurrent.

What do these statements mean?

The definition means that the price is called “recurrent” on a given timescale h if it retraces within that timescale, in the past or future. “Retracing in the past” is equivalent to saying that the price now is retracing a potential transient zone established before.

The proposition implies that the price is almost everywhere (“almost surely”) recurrent. It does that by construction: it says that given the price it is possible to find a (large-enough) timescale h and a (sufficiently small) interval bound k around the price so that every price within the interval is recurrent.

In the recurrence statistic indicator we use a simple corollary of the Proposition 1:

Corollary 1: For a given positive integer price h , there exist a set of integers k_i and a set of times t_i relative to Timeframe T for which the price $X_T(t_i)$ will *not* be $h(T)$ recurrent (i.e. will be transient) in the interval $[X_T(t_i)-k/2, X_T(t_i)+k/2]$.

The corollary relies on the premise that we typically choose the value of the h , so with this choice of a finite (i.e. *small*) number h , we will surely find some areas that will not be returned to. The interval now means a different thing than in the original proposition, it is now the interval in which the price will *not* be returned to, so it is a transient interval! The Recurrence statistic indicator provides us with various statistical quantities, notably the frequency of true transient zones (not retraced for a chosen h), their probability of resolution (i.e. turning a potential transient zone into a recurrent zone by price retrace within h bars) and the average width of the transient interval k .

A simple way of using this indicator is to play with the numbers for h to find the frequency of true transient zones and more importantly to obtain the clearance (retracing) frequency as large as possible while still maintaining a reasonable timeframes for trading. The expectation is that the historical frequencies obtained from the data, when calculated over a large period of time, will change little for another period of time, so they can be used like probabilities.

This approach is in the true spirit of the Similarity thread, one searches for historically rare anomalies in the making, and then trades with the expectation that the anomaly will be resolved. With this brilliant contribution from Eurussd it is possible to have true high-probability trades. The main input from Recurrence Statistic is that it quantifies the corresponding probabilities.

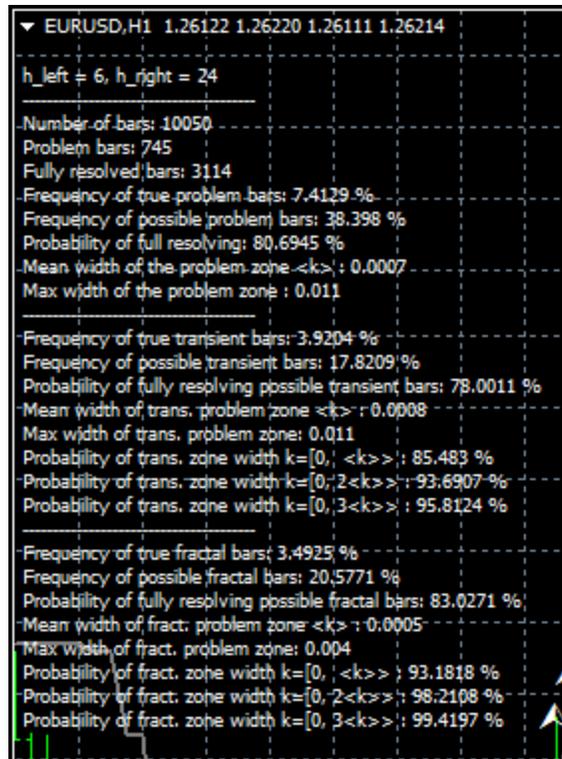


Fig. 1. A typical statistics output from Recurrence Statistic V4. First the input parameters h_{left} and h_{right} are given, then the total statistics for all “problem bars”, i.e. the potential transient zones, and finally separate stats for mid-bar “transient” and end-of-the-bar “fractal” candles is provided.

Typical basic statistics output:

Number of bars: total number of bars over which the statistics was calculated

Problem bars: number of true transient zones – we typically want as little as possible, while still having manageable trading times

Fully resolved bars: number of potentially transient areas that turned out to be recurrent.

Frequency of true problem bars: Problem bars divided by Number of bars. This quantity should be as little as possible, since in the simple strategy, one will not achieve “take profit” if the zone is not cleared.

Frequency of possible problem bars: Calculated as Problem bars plus Fully resolved bars, this sum divided by Number of Bars. This is an important quantity

as it limits the number of potential trades based on the indicator. Clearly, we would prefer to have as many as possible of possible trades with most of them ending up in profit, so there is a compromise one needs to make with respect to trade frequency and probability of success.

Probability of full resolving: Calculated as Frequency of possible problem bars minus the Frequency of true problem bars, divided by Frequency of possible problem bars. **The most important quantity!** This means exactly what it says, if you have a transient zone in the making (the left-transient zone), this tells you the probability of it becoming recurrent. Again note that we used the relationship frequency -> probability. Frequency is something we calculate from the historical data, probability is something that we imply for the future behavior.

All these quantities are calculated separately for the fractal and the transient case.

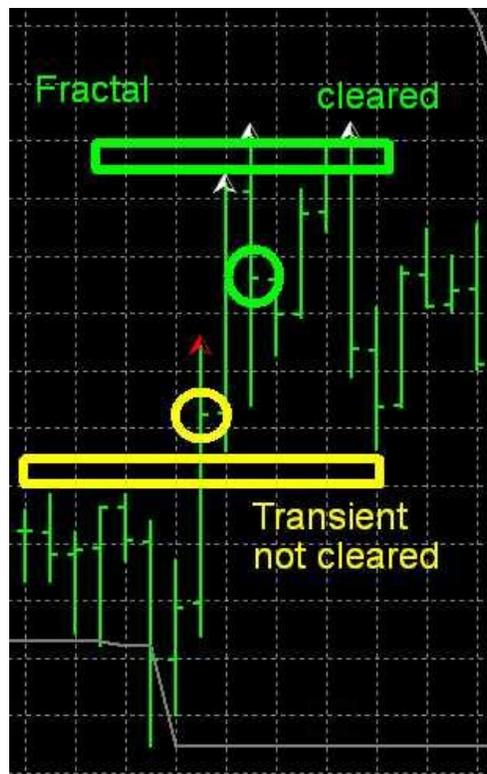


Fig. 2. Difference between the “fractal” (end-of-the-bar) and “transient” (mid-bar) zones. In this example the transient (yellow) is not cleared – it has not become recurrent, while the fractal one (green) did. The indicator paints arrows – by default White are cleared, while DodgerBlue and Red denote uncleared, “true” zones.

Fractal and transient zones:

The indicator differentiates between the fractal and transient zones and does the statistics separately. Fractal zones are non-recurrent zones at the end of the candle, because true fractal zones (typically with $h=2$) at the end of the candle are known as Bill Williams fractals. Transient zones are the mid-candle ones. So, if your zone of interest is a bar-end or a fractal zone, you can look at the quantities for fractal zones (for example the “Probability of fully resolving possible fractal bars”), whereas if you are wondering what the odds are for clearing a mid-bar or transient zone, the quantity to look for is the “Probability of fully resolving possible transient bars”.

The k business:

The indicator, as defined in the corollary, calculates the average width of the unresolved problem bars. It is a useful quantity, since it may provide an even higher probability trades with slightly more modest goals.

The price does not necessarily clear the potentially transient zone. What often happens is that it comes close to retracing/clearing it. This essentially means that a narrow “true” transient zone remained. How narrow? The indicator tries to answer this question by giving an average width $\langle k \rangle$ of the zone (as always separately for end-of-bar=fractal and mid-bar=transient scenarios). Additionally, it gives the probability that the price will retrace within $1 \cdot \langle k \rangle$, $2 \cdot \langle k \rangle$, $3 \cdot \langle k \rangle$.

For example, in Fig 1 we see that the $\langle k \rangle$ for the end-of-the bar with specified input parameters is $0.0005 = 5$ pips. While the probability that the zone will be cleared within specified h_{right} is 83%, the probability that the price will reach within 15 pips ($3 \cdot \langle k \rangle = 3 \cdot 5$ pips) to clearing it is 99.4197% !

The arrows:

The indicator plots arrows (See Fig 2) on the candles where there are either potential transient zones or fully resolved zones (default in white color) or true transient zones (default in dodgerblue for down and red for up). The up and down arrow mean that the zone is towards up or down.

kprsa, 16.11.2014