INTRODUCTION

Chemical synapses are highly specialized structures that enable neurons to exchange signals, or to send signals to non-neural cells such as muscle fibers. Even though there is a staggering diversity of synapse morphologies and types in the brain, the fundamental process of synaptic transmission is always the same. A presynaptic membrane potential depolarization, typically caused by the arrival of an action potential, triggers the release of neurotransmitter, which then binds to receptors that, in turn, generate a response in the postsynaptic neuron.

A key quantity in neural circuits is the synaptic efficacy or strength, which varies over time. Cellular processes such as long-term potentiation and depression contribute to the patterning of the nervous system during development, and are thought to constitute the basis of learning and memory (Morris, 2003). Slow and long-lasting homeostatic processes adjust synaptic strength to maintain circuit activity within functional regimes (Turrigiano and Nelson, 2004). In addition, a whole range of activity-dependent processes exist that modulate synaptic efficacy continuously on very short time scales ranging from milliseconds to minutes (for reviews, see Zucker and Regehr, 2002; Fioravante and Regehr, 2011). Unlike long-term and homeostatic plasticity, short-term plasticity, the topic of this review, has a direct influence on the computation performed by neural circuits as these dynamics take place on the time scale of stimulus-driven activity, neural computations and behavior.

Broadly, short-term plasticity can be classified as synaptic depression and facilitation. Depression refers to the progressive reduction of the postsynaptic response during repetitive presynaptic activity, while facilitation is an increase synaptic efficacy. Each of these may be caused by a range of different mechanisms with different time constants, and the two forms are not mutually exclusive. For instance, a particularly well-studied example of a strongly depressing synapse is the calyx of Held, a giant synaptic terminal in the mammalian auditory brainstem (Schneggenburger and Forsythe, 2006). A closer look at the underlying mechanisms, however, reveals that the response is also modulated by facilitation, which is however, partially masked by depression. In fact, most synapses express some combination of these two mechanisms, but with considerable variability between different neuron types (Wang et al., 2006).

The purpose of this review is to summaries models of short term plasticity, to discuss their biological background and plausibility, and to provide a guide for selecting an appropriate model and level of detail. The focus here is on the mechanistic aspects of these models, for a review of functional implications see Abbott and Regehr (2004). The review begins with a reminder of the main processes involved in synaptic transmission. Next, the vesicle depletion model and its variants will be introduced as a canonical model for short term plasticity. Finally, several additions to this class of models will be discussed that were required to explain more recent experimental findings.

PRINCIPLES OF SYNAPTIC TRANSMISSION

Almost all factors contributing to short term plasticity are located in the presynaptic terminal. To identify the relevant variables required in models, we begin with a brief review of the main events following the arrival of a presynaptic action potential at a synapse, as illustrated in Figure 1. The site where synaptic transmission of neural activity is initiated is called the active zone (AZ), a presynaptic morphological specialization where vesicles containing neurotransmitter and proteins required for the release process are clustered. The AZ is opposed by the postsynaptic density (PSD), an area that contains a large number of different proteins implicated in synapse maintenance and plasticity. In addition to a whole variety of structural and signaling complexes, the PSD contains the bulk of the neurotransmitter receptors mediating the postsynaptic response.

Neurotransmitter release from vesicles located at the AZ is initiated by an elevation of the intracellular calcium concentration \([\text{Ca}^{2+}]\), due to opening of voltage gated calcium channels.
Hence the second variable required in a synapse model is the number of vesicles \( N(t) \) available for release. Again, as will be discussed in more detail below, the number of release-ready vesicles changes over time since the occupancy of the pool changes during neural activity. Vesicle number and release probability are the key ingredients for a model of presynaptic transmitter release:

\[
T(t) = p(t) \cdot N(t) \tag{1}
\]

Here \( T(t) \) is the amount of transmitter released into the synaptic cleft at time \( t \). Simulating a highly realistic synapse model using this expression would require a precise, time continuous model of calcium influx and vesicle cycling. However, since the release probability dramatically increases upon the arrival of a presynaptic action potential from a resting value of almost zero, it is usually sufficient to update these quantities only once every time a presynaptic action potential arrives.

Finally, the released transmitter diffuses through the synaptic cleft and binds to receptors to generate a postsynaptic response, the main quantity of interest in synapse models. Here, we focus on the action of ionotropic receptors, which contain an ion channel that opens when transmitter is bound. The kinetics of such a response is determined by the rates of transmitter binding and unbinding and opening and closing of the channel, as well as transitions to and from desensitized states. The simplest model of this process is when the postsynaptic conductance is proportional to the amount of transmitter released:

\[
g(t) = g_m T(t) \tag{2}
\]

The peak conductance is denoted by \( g_m \). If the time course of the response is relevant, for instance to distinguish between fast AMPA receptor and slow NMDA receptor mediated transmission, alpha functions, double exponential models, or simple kinetic models are useful to model this process (Destexhe et al., 1994b; Roth and Rossom, 2009).

Numerous studies have been devoted to assessing the release probability and quantal content of synapses in various brain areas and neurons types. As will be shown below, this is generally achieved through model-based analysis, which is possible because the synapse models provide a good mapping between experimental observables, usually the postsynaptic current and its variance, and the underlying synaptic parameters. A comprehensive overview of parameters of a range of neuron types assessed in this way can be found in a review by Branco and Staras (2009).

**THE VESICLE DEPLETION MODEL AND EXTENSIONS**

**VESICLE DEPLETION AS MAIN CAUSE OF SYNAPTIC DEPRESSION**

The outline in the preceding section hints that presynaptic vesicles are a limited resource, and that their depletion during ongoing activity can lead to a suppression of the postsynaptic response. The first formal model of such a process was published by Liley and North (1953), even before synaptic vesicles were discovered by De Robertis and Bennett (1955). It sought to explain synaptic depression during brief tetanic stimulation of the rat neuromuscular junction, and was based on...
of short term plasticity

Models of short term plasticity

the assumption that releasable neurotransmitter is produced at a limited rate. Tetanic stimulation was assumed to cause transmitter depletion and a concomitant reduction in postsynaptic response. This process is described by a simple first order kinetic model:

\[
\frac{dn(t)}{dt} = \frac{1 - n(t)}{\tau_r} - \sum_j \delta(t - t_j) \cdot p \cdot n(t)
\]

where \(n(t)\) is the occupancy of the release pool, bounded between zero and one, \(\tau_r\) the time constant of the vesicle replenishment, and \(t_j\) the presynaptic spike times. Note that in this and all following equations, the dynamic quantity, here \(n(t)\), is evaluated before the delta function [as in \(n(t - \epsilon)\), here the \(\epsilon\) is omitted for clarity]. The release term reduces the vesicle pool occupancy by \(T(t) = \dot{p} \cdot n(t)\), which is proportional to the postsynaptic response (see Equation 2). Experiments suggest that the recovery time constant is typically in the order of seconds. Equation (3) describes a continuous form of the model, which may be inappropriate for synapses with a small number of releasable vesicles, as it is often the case. Then a discrete form should be used where the release pool occupancy \(n(t)\) is replaced by the vesicle number \(N(t)\). In this case, a discrete form is also required to accurately model the stochasticity of synapses.

This model predicts an exponential decay of the postsynaptic response during stimulation at a constant rate, and an inverse relation between input frequency \(v\) and steady state level of depression \(n_{\infty} = 1/(pv\tau_r + 1)\) (Figures 2A,E). It was found to fit responses recorded from some depressing synapses very well (Liley and North, 1953; Tsodyks and Markram, 1997), including EPSCs during stimulation of the calyx of Held with in vivo-like activity patterns (Hermann et al., 2009). However, often synapses show substantial deviations. In particular, the steady state values decrease more slowly with increasing frequency than the inverse behavior predicted here.

**SYNAPTIC FACILITATION**

To explain such deviations from the deletion model, it was first suggested by Betz (1970) to extend it by release probability facilitation that counteracts depression. Potential underlying mechanism of facilitation are an accumulation of residual calcium in the synaptic terminal (Atluri and Regehr, 1996; Blatow et al., 2003; Felmy et al., 2003), which causes rapid VGCC facilitation (Katz and Miledi, 1968; Borst and Sakmann, 1998; Cuttine et al., 1998; Mochida et al., 2008). A simple phenomenological model of such processes is to increase the release probability after each presynaptic spike (Betz, 1970; Varela et al., 1997; Markram et al., 1998):

\[
\frac{dp(t)}{dt} = \frac{p_0 - p(t)}{\tau_f} + \sum_j \delta(t - t_j) \cdot a_f \cdot (1 - p(t))
\]

Here \(p_0\) is the baseline release probability, \(a_f\) the amount of facilitation per action potential and \(\tau_f\) the recovery time constant. The time constant is typically in the range of tens of milliseconds, much faster than vesicle replenishment. Therefore, facilitation is usually observed during more intense periods of activity. Steady-state facilitation approaches \(p_{\infty} = (p_0 + v a_f \tau_f)/(1 + v a_f \tau_f)\) for a stimulus with constant frequency \(v\) (Figure 2E).

The net effect of the combined model of facilitation and vesicle depletion depends strongly on the basal release probability: for a small \(p_0\), facilitation can have a substantial effect since it is not masked by rapid vesicle pool depletion, and for large values depression will dominate over depletion (Figure 2B). As a general rule, it appears that synapses with a larger vesicle pool also tend to have a higher release probability (Dobrunz and Stevens, 1997). Hence facilitation is expected to be more dominant at “smaller” synapses.

This extension of the depletion model can account quite well for data where the simpler depletion model fails, in particular the relationship between stimulus frequency and steady-state response amplitude (Varela et al., 1997; Markram et al., 1998). For instance, a comprehensive survey of cells in the medial prefrontal cortex has shown that this model can fit a wide range of different behaviors encountered in such data sets, despite large variability in the relative contribution of depression and facilitation (Wang et al., 2006).

This depletion model with facilitation has become very popular as a canonical model for short term plasticity. It has, either in the form given here (Equations 3, 4) or using a slightly different set of equations as introduced by Tsodyks et al. (1998), been used in many studies investigating the functional importance of short term plasticity (see e.g., Abbott et al., 1997; Tsodyks et al., 1998, Fuhrmann et al., 2002; Mongillo et al., 2008; Pfister et al., 2010). As usual, however, a closer experimental investigation of synapses has shown that this relatively simple and intuitive model lacks potentially important detail, as will be discussed in the following sections.

**USE-DEPENDENT VESICLE REPLACEMENT**

An important observation at odds with the deletion model is that vesicle replenishment can accelerate after intensive stimulation. This effect was found to depend on an increase in intracellular calcium concentration, and to occur in a physiological range of input firing rates (Dittman and Regehr, 1998; Stevens and Wesseleng, 1998; Wang and Kaczmarek, 1998; Sakaba and Neher, 2001; Fuhrmann et al., 2004; Hosoi et al., 2007). Enhanced vesicle replenishment can be included in the depletion model by adding some form of activity-dependent component to Equation (3). Two slightly different approaches have been proposed, both capable of explaining the slow reduction in steady state depression for strong stimuli that the simple deletion model fails to replicate.

The first model, introduced by Fuhrmann et al. (2004) to reproduce depression at cortical synapses, was based on the idea that presynaptic activity directly modulates the time constant \(\tau_r\) of vesicle replenishment in Equation (3) above:

\[
\frac{d\tau_r(t)}{dt} = \frac{\tau_{\tau_0} - \tau_r(t)}{\tau_{FDR}} - a_{FDR} \tau_r(t) \cdot \sum_j \delta(t - t_j)
\]

Here each presynaptic action potential reduces the time constant by \(a_{FDR} \tau_r(t)\), which recovers to its resting value \(\tau_{\tau_0}\) with
Hennig Models of short term plasticity

FIGURE 2 | Summary of the key characteristics of the models discussed in this review. (A–D) Postsynaptic response for the different models during stimulation at different frequencies. (A) The vesicle depletion model (Equation 3) predicts exponential decay of the response and an inverse relation between stimulus frequency and steady-state amplitude. A higher release probability causes faster and stronger depression [compare upper and lower graph, see also panel (E)]. (B) The depletion model with facilitation (Equations 3, 4) predict a transient response increase during high-frequency stimulation. For a low basal release probability \( p_0 \) the response remains elevated (top graph), while for higher \( p_0 \) vesicle depletion masks facilitation [bottom graph, see also panel (E)]. (C) Use-dependent vesicle replenishment (Equation 6) increases the steady-state response. (D) As panel (C), but with added slow use-dependent suppression of release probability. Here the postsynaptic response continues to slowly decay when the depletion model reaches steady-state [compare (C) and (D)]. (E) Steady-state response magnitude as a function of input frequency for the depletion model (circles) and the depletion model with facilitation (dashed lines). (F) Same as (E), but for the depletion model with use-dependent replenishment (UDE, circles) and the UDE model with slow suppression of release probability (RS, dashed). Note that the latter increases depression in particular at low frequencies. (G) Occupancy of the releasable vesicle pool for the models in panel (F). It is less depleted for the RS model as steady-state depression is mediated by the reduction in release probability. Parameters: \( \tau_r = 1 \text{s} \), \( a_f = 0.3 \), \( \tau_f = 0.2 \text{s} \) [no facilitation in (C,D)], \( a_e = 0.4 \), \( \tau_e = 0.1 \text{s} \), \( a_i = 0.01 \), \( \tau_i = 10 \text{s} \).

In these experiments, the recovery time course was fit by two exponential functions, suggesting the combined action of at least two processes. This can be modeled by augmenting a constant background replenishment with a low rate \( (k_r = \frac{1}{\tau_r}) \) with an activity-dependent component:

\[
\frac{dk_e(t)}{dt} = -\frac{k_e(t)}{\tau_e} + a_e \cdot \sum_j \delta(t - t_j) \cdot (1 - k_e(t)) \quad (6)
\]

This process is activated by presynaptic activity, leads to an increment \( a_e \) of the replenishment rate for each action potential, a time constant \( \tau_{DR} \) in the order of hundreds of milliseconds. A very similar model with a non-linear relation between intracellular calcium concentration and recovery rate was proposed to explain the different kinetics observed at hippocampal and cerebellar synapses (Dittman and Regehr, 1998; Dittman et al., 2000).

Alternatively, it may be assumed activity leads to a temporary enhancement of vesicle replenishment. This is based on the observation that high-frequency stimulation causes a fast but short-lived component of recovery from depression, which is absent after weaker stimulation (Wang and Kaczmarek, 1998).
and decays with a time constant $\tau_r$ in the range of 10–100 ms. Equation (3) then becomes:

$$\frac{dn(t)}{dt} = \left( k_r + \bar{k}_r e(t) \right) (1 - n(t)) - \sum_j b(t - t_j) \cdot p(t) \cdot n(t)$$

(7)

where $\bar{k}_r$ is the peak rate of activity-dependent vesicle replenishment. This model predicts weaker steady-state depression at high frequencies (Figures 2C,F), and has been shown to rather accurately reproduce the vesicle pool kinetics (Hosoi et al., 2007) and steady-state behavior at the calyx of Held (Wong et al., 2003; Hennig et al., 2008).

The biophysical mechanism behind use-dependent vesicle replenishment is still not well understood. It appears clear that it depends on calcium influx (Wang and Kaczmarek, 1998; Sakaba and Neher, 2001; Hosoi et al., 2007), but it has been difficult to experimentally disentangle the role of calcium-dependent vesicle recruitment and calcium-dependent endocytosis, perhaps because most studies so far used extremely strong and unphysiological stimuli to deplete the vesicle pool. A recent study suggests that these two processes may in fact be linked, and that perhaps the speed at which release sites are made available by endocytosis is an important rate limiting step during high frequency transmission (Yao and Sakaba, 2012). Use-dependent replenishment may then reflect faster recruitment due to more efficient endocytosis.

A main function of this mechanism appears to maintain the ability of a synapse to transmit during sustained periods of high activity (Wong et al., 2003; Hosoi et al., 2007). It is as such an important, and often overlooked component of short term plasticity that has implications for transmission of varying firing rates. In addition, it has been suggested to improve transmission by broadening the range over which information about rate and rate changes are reliably transmitted (Fuhrmann et al., 2004; Yang et al., 2009). Which of the two models discussed here is more appropriate is unclear. The difference between the two models is that enhanced replenishment is unbounded in Equation (5), but bounded in Equation (6). Hence the former predicts a faster decrease of the steady state response amplitude with increasing frequency, which more quickly settles to a constant value. It is therefore possible that it underestimates the amount of depression at some synapses, but this would require a more exhaustive comparison with data.

**SLOW MODULATION OF RELEASE PROBABILITY**

A further omission of the depletion model is that activity-dependent release probability suppression may also contribute to synaptic depression (Xu and Wu, 2005; Mochida et al., 2008). Potential mechanisms include VGCC inactivation (Forsythe et al., 1998; Patil et al., 1998) or activation of presynaptic autoreceptors such as mGluRs or AMPARs, which in turn can cause a reduction of the release probability (Takahashi et al., 1996; Takago et al., 2005). A possible molecular route of such effects is calcium/calmodulin (Lee et al., 1999). Postsynaptic release of endocannabinoids has also been shown to suppress synaptic strength over short time scales, but the mechanisms are currently not well understood (Brenowitz and Regehr, 2005). Overall, the degree to which these mechanisms are relevant under physiological conditions is still not fully understood. For instance, release probability suppression has been reported to strongly contribute to synaptic depression during weak activity at the calyx of Held (Xu and Wu, 2005), but this effect may be more pronounced at immature synapses were morphological development renders synaptic transmission less effective (Renden et al., 2005; Nakamura et al., 2008).

A generic model incorporating both release probability facilitation and depression can be constructed by extending Equation (4) by an activity-dependent modulation of the baseline release probability $p_0$ (Billups et al., 2005; Hennig et al., 2008):

$$\frac{dp_0(t)}{dt} = -\frac{p_0 - p_0(t)}{\tau_i} - \sum_j b(t - t_j) \cdot a_i \cdot p_0(t)$$

(8)

Here the baseline release probability $p_0(t)$ is reduced by a constant fraction $a_i$ after each spike, and recovers back to $p_0$ with a time constant $\tau_i$ in the order of several seconds. Then depression of release probability is proportional to the incoming spike rate. An alternative form, which models the activation of autoreceptors, is to replace the term on the right-hand side with $\sum_j b(t - t_j) \cdot a_i \cdot p_0(t) \cdot p(t) \cdot n(t)$. In this case, depression of release probability is release-dependent. Combinations of both mechanisms are also possible, as shown by Hennig et al. (2008). In combination with the depletion model and facilitation (Equations 3 or 6, and Equation 4), this model can account for a slow form of depression that follows initial rapid vesicle depletion (Figures 2D,F), as observed at GABAergic synapses (Kraushaar and Jonas, 2000) or the calyx of Held (Hennig et al., 2008) during prolonged stimulation.

The analysis of the steady-state behavior the model reveals an interesting further property (Hennig et al., 2007). If the release probability is assumed to vary slowly compared to the effective vesicle replenishment rate $\bar{k}_r$, the quasi-stationary solution of Equation (3) with use-dependent vesicle replenishment (Equation 6) is $n_\infty p_c = \bar{k}_r (1 - n_\infty)$, where the index $c$ indicates that $p_c$ is constant over the time interval considered, and we obtain $n_\infty = k_c/(p_c + k_c)$. This solution is valid when all fast processes (e.g., facilitation) have settled to their stationary values. If the release probability is now changed by a small amount to $p_c = \alpha p_c$, then the vesicle pool occupancy settles to a value that differs by a factor of $n_\infty / n_\infty = (p_c + k_c)/(\alpha p_c + k_c)$.

Hence a slow reduction in release probability will not only slowly depress the postsynaptic response, but also increase the size of the releasable vesicle pool (Figure 2G). This corresponds to a transfer of depression from vesicle depletion to a reduction of release probability. The net effect is a decrease in postsynaptic response that is slower than the change in release probability, and a concomitant refilling of the vesicle pool. Analysis of synaptic depression at the calyx of Held during prolonged stimulation support this conclusion, and suggest that it is, in part, mediated by mGluR autoreceptor activation (Billups et al., 2005; Hennig et al., 2008).
A CLOSER LOOK AT RELEASE PROBABILITY

A central variable in the models discussed is the release probability, and so far the effect of activity was assumed to be linear. This is however, incompatible with the steep non-linearity that couples presynaptic calcium influx to release rate (Bollmann et al., 2000; Schneggenburger and Neher, 2000; Lou et al., 2005). If we assume that the effects of facilitation and depression discussed above such as accumulation of residual calcium, channel facilitation or inactivation, have a linear effect on the calcium concentration, this non-linearity would predict a far more drastic effect on the release rate. In fact, early studies already found that a third to fourth-power relationship is a better model for facilitation than a linear model (Zengel and Magleby, 1982).

An analysis of synaptic depression at the calyx of Held by Xu and Wu (2005) further confirms this intuition. This study suggested that depression during slow stimulation (in the range between 1 and 10 Hz) is primarily mediated by a reduction in calcium influx, while vesicle depletion is only effective at higher frequencies. Interestingly the model presented in the preceding section qualitatively reproduces this effect. As shown in Figure 2F, slow depression of release probability has a significant effect at low frequencies when compared to an equivalent depletion model, which becomes weaker with increasing frequency. However, as shown above this model also predicts that the depression at higher frequencies is still due to reduced release probability, which replaces vesicle depletion during sustained activity. There is some experimental evidence based on fluctuation analysis in support of this hypothesis (Hennig et al., 2008), but it will be interesting to see if alternative vesicle depletion models can also account for these findings.

AUGMENTATION AND POST-TETANIC POTENTIATION

Augmentation and post-tetanic potentiation are two slowly developing and long-lasting forms of synaptic enhancement (Fisher, 1997; Zucker and Regehr, 2002). They are induced by prolonged stimulation of the synapse, and vary in their activation and relaxation kinetics. The faster form, with time constants of seconds, is typically referred to as augmentation, whereas post-tetanic potentiation operates on time scales of tens of seconds. It appears that these processes are caused by an increase in release probability, which can be occluded by depression due to vesicle depletion during ongoing stimulation (Habets and Borst, 2007). While early studies proposed accumulation of residual calcium at the synaptic terminal as a primary mechanism (Zengel and Magleby, 1982; Zucker and Lara-Estrella, 1983; Habets and Borst, 2007), more recent work also implicated PKC activation (Korogod et al., 2007) or calmodulin/CaM kinase II activity (Fiumara et al., 2007).

A model which could account for a range of findings in data from the frog and toad neuromuscular junction was proposed by Zengel and Magleby (1982). They proposed that facilitation (F), augmentation (A) and post-tetanic potentiation (P) affect release probability in a multiplicative manner:

\[ P(t) \propto F(t)A(t)P(t) \]  

Each process follows first order kinetics, and facilitation was best captured by including a fast and a slow component (see also Zucker and Lara-Estrella, 1983). While facilitation required a fourth-power relationship between the corresponding state variables and release rate, it was sufficient to assume a linear dependence for augmentation and potentiation. This points to different potential sites of action of these mechanisms as outlined above. In addition, it was found that augmentation increases with longer stimuli. This was modeled by including a time-dependent increase in activation rate \( a = a_0 e^{-T} \) (where \( v \) is the stimulus frequency, \( T \) is stimulus time and \( z \) a constant), but could also indicate the presence of multiple first-order processes acting on different time scales (Drew and Abbott, 2006; Hennig et al., 2008). For instance, activation of presynaptic NMDA receptors has also been shown to enhance release probability, with a time course in the order of minutes (Duguid and Smart, 2004).

So far, few theoretical studies have investigated the implications of slow enhancement of release using detailed models. A simple, phenomenological model based on Equation (4) above, where time constants were chosen in the range of augmentation, suggests a potential role in short term memory (Mongillo et al., 2008).

THE OTHER SIDE: RECEPTOR DESSENSITIZATION

The time course of the postsynaptic response depends not only on the amount of released transmitter and its time course, but also on the kinetics of the receptors. The interplay of these factors with synapse morphology has been investigated in great detail with Monte Carlo simulations (Stiles et al., 1996; Franks et al., 2003; Coggan et al., 2005; Postlethwaite et al., 2007), which are in particular useful to understand the sources of variability at synapses. The semi-quantitative models discussed in this review cannot easily accommodate this level of detail, but can still be extended to include salient aspects of the postsynaptic response (Destexhe et al., 1994a; Roth and Rossum, 2009).

Apart from the response latency and duration, desensitization is an important property of receptors which has been shown to contribute to synaptic depression during physiological activity levels (Trussell et al., 1988, 1993; Jones and Westbrook, 1996; Neher and Sakaba, 2001). A simple but effective approximation of the state of the population of receptors \( D(t) \), can be modeled using first order kinetics:

\[
\frac{dD(t)}{dt} = \frac{1 - D(t)}{\tau_D} - \sum_j \delta(t - t_j) \cdot a_j \cdot p(t) \cdot n(t) \cdot D(t) 
\]

The quantity \( D(t) \) represents the fraction of non-desensitized receptors. Recovery from desensitization \( \tau_D \) is typically in the order of tens of milliseconds, such that it is only effective during intense episodes of activity. The postsynaptic response is then expressed as \( R(t) = g_m D(t) \cdot n(t) \cdot p(t) \), where \( g_m \) is the peak conductance.

This basic model captures synaptic depression due to desensitization well. In particular, simulations have shown that a main effect is the masking of presynaptic facilitation at high stimulus frequencies (Jones and Westbrook, 1996; Wong et al., 2003). Yet in this form the model obviously neglects the time course of the postsynaptic potential, which can also be affected by...
STOCHASTICITY OF SYNAPSES

Transmitter release is a stochastic process, and as a consequence the magnitude of the postsynaptic current evoked by each presynaptic action potential fluctuates from time to time. Due to the quantal nature of synaptic transmission, the variance of the postsynaptic response is described by binomial statistics, with a predicted variance of $\text{Var}(g(t)) = g_m \cdot N(t) \cdot p(t) \cdot (1 - p(t))$ (Del Castillo and Katz, 1954). This shows that changes in the synaptic parameters due to short term plasticity will not only cause changes in the average postsynaptic response, but also in the magnitude of the fluctuations, as measured by the coefficient of variation:

$$CV(g(t)) = \sqrt{\frac{1 - p(t)}{N(t) p(t) g_m}}$$ (11)

This value is high when the release probability or the number of release-ready vesicles is small, as, for instance, often found for cortical neurons (Wang et al., 2006; Brémaud et al., 2007). The expression also shows that stochastic effects are bound to be more important when synaptic depression is dominated by vesicle depletion. In addition, the entire vesicle cycle, which includes vesicle replenishment, consists of stochastic events. In contrast, the influx of calcium during an action potential, which triggers transmitter release, is considered a much more salient event, and is therefore expected to contribute much less to postsynaptic response variability. To model the main sources of stochasticity of synaptic, the models discussed above can be directly cast into a stochastic form by simulating vesicle release and replenishment as random events (see e.g., de la Rocha and Parga, 2005; Yang et al., 2009; Rosenbaum et al., 2012).

Stochastic models are extremely useful for quantitative evaluation of models of synaptic transmission and plasticity, since postulated changes in $N$ and $p$ have predictable effects on variability that can be directly tested experimentally (see e.g., Quastel, 1997; Scheuss and Neher, 2001; Brémaud et al., 2007). This type of analysis requires a careful dissection of synaptic function, since for instance conductance changes through receptor desensitization may be mistaken for presynaptic effects if not properly controlled for. Recent work exploiting this approach provided evidence for substantial variability of synaptic parameters for synapses between cortical pyramidal neurons, the presence of multi-quantal release (Loebel et al., 2009), and the coordination of pre-synaptic release probability and postsynaptic synaptic strength (Hardingham et al., 2010).

Studies investigating stochastic synapse models have reported several effects indicating that this also has important implications for neural computations and network function. For instance, shot noise due to stochastic release can increase the output firing rate of a neuron operating in a fluctuation driven regime when compared to deterministic dynamics (de la Rocha and Parga, 2005). The same study also showed that short term depression in stochastic synapses causes a further, non-monotonic modulation of output firing in presence of input correlations (see also Rosenbaum et al., 2012). Such effects were further analyzed by Rosenbaum et al. (2013), who showed that, unlike for a deterministic model, a stochastic synapse with short term depression can significantly de-correlate neural activity. Finally, an analysis of stochastic models including slow release probability modulation and activity-dependent vesicle replenishment suggested that multiple mechanisms of short term plasticity may act synergistically to maintain stable information transmission over a broad range of input frequencies (Yang et al., 2009). Overall, however, the models used so far to analyze stochastic effects were mostly rather simple, typically only the depletion model was considered, and assumed constant random inputs to the neuron (see Merkel and Lindner, 2010, for an extension).

OUTLOOK

Theoretical models have contributed much to our understanding of synaptic transmission and short term plasticity by providing a framework to express conceptual models in rigorous terms, and to derive quantitatively testable hypotheses. The models discussed here capture the central biophysical processes involved in synaptic transmission in relatively simple mathematical form, such that an exact or at least approximate analytical treatment is possible. Moreover, key variables in these models have direct measurable correlates. This supports analysis and comparison with data, as often exploited for deriving synaptic parameters from experimentally recorded synaptic currents. It is however not straightforward to experimentally interfere with short term plasticity in intact neural circuits in a targeted manner, for instance to assess functional implications and consequences. Therefore, these models are also a valuable tool that enables analysis beyond the experimentally feasible.

The basic depletion model with facilitation has passed the test of time, which nicely illustrates the success of simple, mathematically tractable phenomenological models in biology. However, as shown here, short term plasticity can be more complicated. In particular slow forms of synaptic depression and facilitation merit more thorough investigation, both in terms of mechanisms and their relevance for neural computations. While the depletion model can very successfully replicate even synaptic responses during in vivo-like activity patterns (Hermann et al., 2009),
slow synaptic modulation may have important effects during firing rate modulation on time scales of tens of seconds (see e.g., Mongillo et al., 2008). A combination of slow facilitation and depression has also been shown to support differential responses to time varying stimuli (Barak and Tsodyks, 2007). These studies show that these mechanisms certainly warrant further investigation.

As shown in this review, even the extended and more complete models of short term plasticity have a relatively simple mathematical form, which will greatly facilitate the understanding of their effects in networks. Perhaps a central question in this context is how far the different mechanisms discussed here have direct functional implications, or rather reflect the biophysical properties and limitations of chemical signaling between neurons. Some of the studies touched upon above and in the previous section suggest the former may be the case (for a more detailed discussion, see e.g., Abbott and Regehr, 2004). On the other hand it is equally plausible some aspects of short term plasticity may related to homeostatic effects or metabolic efficacy of synapses, issues that have received little attention so far and are now easily testable in models. Addressing such questions may require the analysis of the models under more physiologically relevant conditions. For example, recent experiments indicate that unreliable synapses with short term plasticity are particularly suited to transmit information contained in brief bursts of activity typically observed in hippocampus (Rotman et al., 2011). Therefore, modeling studies specifically investigating synapses in their “natural habitat” of recurrent networks should allow us to refine and consolidate such hypotheses, and to establish more of the much sought-after links between neural biophysics and brain function and dysfunction.

ACKNOWLEDGMENTS
I thank Mark van Rossum for comments. This work was funded by an MRC Career Development Award (G0900425).

REFERENCES
Betz, W. J. (1970). Depression of transmission, see e.g., Abbott and Regehr, 2004. On the other hand it is equally plausible some aspects of short term plasticity may related to homeostatic effects or metabolic efficacy of synapses, issues that have received little attention so far and are now easily testable in models. Addressing such questions may require the analysis of the models under more physiologically relevant conditions. For example, recent experiments indicate that unreliable synapses with short term plasticity are particularly suited to transmit information contained in brief bursts of activity typically observed in hippocampus (Rotman et al., 2011). Therefore, modeling studies specifically investigating synapses in their “natural habitat” of recurrent networks should allow us to refine and consolidate such hypotheses, and to establish more of the much sought-after links between neural biophysics and brain function and dysfunction.

ACKNOWLEDGMENTS
I thank Mark van Rossum for comments. This work was funded by an MRC Career Development Award (G0900425).

REFERENCES


Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.