Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells

Hateren, J.H. van; Snippe, H.P.

Published in:
Vision Research

PERGAMON

Vision Research 41 (2001) 1851–1865

Vision Research

www.elsevier.com/locate/visres

Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells

J.H. van Hateren *, H.P. Snippe

Department of Neurobiophysics, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands

Received 17 August 2000; received in revised form 12 January 2001

Abstract

Models are developed and evaluated that are able to describe the response of blowfly photoreceptor cells to natural time series of intensities. Evaluation of the models is performed using an information theoretical technique that evaluates the performance of the models in terms of a coherence function and a derived coherence rate (in bit/s). Performance is gauged against a maximum expected coherence rate determined from the repeatability of the response to the same stimulus. The best model performs close to this maximum performance, and consists of a cascade of two divisive feedback loops followed by a static nonlinearity. The first feedback loop is fast, effectively compressing fast and large transients in the stimulus. The second feedback loop also contains slow components, and is responsible for slow adaptation in the photoreceptor in response to large steps in intensity. Any remaining peaks that would drive the photoreceptor out of its dynamic range are handled by the final compressive nonlinearity.

Keywords: Light adaptation; Natural stimuli; Computational model

1. Introduction

Light intensities in the natural environment of an organism vary considerably. Not only long-term variations in light level are present, as those originating from the cycle of day and night or from the change of seasons, but relatively large variations also occur on a much shorter time scale. This happens for example when moving through different sections of a landscape, such as going from open terrain into the shade of a group of trees. It also happens when the eye shifts its gaze from brightly lit parts of a scene to shaded areas. The photoreceptors of the eye therefore have to cope with variations in light level of at least 2–3 log units in relatively short stretches of time (van Hateren, 1997). This is more than the typical dynamic range of neurons, with dynamic range defined here as the ratio of the maximum response and the noise level. To solve this problem, the photoreceptors of many, if not all, species can quickly adjust their gain to changing light levels. The purpose of the present article is to develop and evaluate models of this gain control in the photoreceptor cells of blowflies, and in particular models that are able to handle naturally occurring series of intensities.

There are at least three reasons why a model of photoreceptor gain control is useful: a physiological, a practical, and a theoretical reason. The first, physiological, reason is that modelling can help to study the physiology of phototransduction and gain control, by pointing to key control loops and suggesting specific experiments. The second, practical, reason for desiring a model of photoreceptor gain control is that it can serve as a preprocessing module for studying visual processing in higher parts of the visual system. More often than not simple, linear processing is assumed for early visual processing. This induces the risk that effects observed at a higher stage are interpreted as properties of that stage, whereas they may in fact be the byproduct of nonlinearities in earlier stages. A full model of higher visual processing needs an adequate preprocessing model, including gain control at the earliest stages. The third, theoretical, reason for wanting a model of gain control is that this will be necessary for understanding the relationship between the properties

* Corresponding author. Tel.: +31-50-3634788; fax: +31-50-3634740.

E-mail address: hateren@phys.rug.nl (J.H. van Hateren).

0042-6989/01/$ - see front matter
PII: S0042-6989(01)00052-9

of natural stimuli and early visual processing. In this approach, visual processing is considered to be optimally adapted to the natural visual environment. Although much progress has been made in this area by the application of information theory (Srinivasan, Laughlin, & Dubs, 1982; Atick & Redlich, 1990; van Hateren, 1992), much of this assumes linear or quasi-linear systems, and a good nonlinear model will be instrumental to extend this approach to a more general range of systems.

At this point in time, there appears to be not yet sufficient information on blowfly phototransduction and light adaptation to produce a detailed physiological model that would work adequately for the type of signals photoreceptors encounter in natural circumstances. Although such a model will ultimately become possible, it is likely to contain many parameters given the complexity of the molecular machinery. For some purposes, a complex model is undesirable. Therefore, we will focus in this article on relatively simple models, with as few parameters as possible. These models will target in particular the second and third purposes mentioned above: the final model will firstly serve as a preprocessing module useful for studies of higher stages of visual processing, and secondly it will help future studies of understanding the information capacity of the early visual system.

For evaluating the various models we use a combination of several recently developed techniques. As a stimulus, Natural Time Series of Intensities (called NTSIs below, van Hateren, 1997) are used, mimicking the statistics of the intensity as a function of time that is normally encountered by individual photoreceptors. Responses from blowfly photoreceptors to this stimulus were measured. From an inversion of the reconstruction method for obtaining information rates (Bialek, Rieke, de Ruyter van Steveninck, & Warland, 1991; Theunissen, Roddey, Stufflebeam, Clague, & Miller, 1996; Haag & Borst, 1997), combined with a nonlinear model, a coherence function of stimulus and response is calculated. This can subsequently be compared with an expected coherence function obtained by looking at response repeatability. As a result, the performance of the models can be evaluated as gauged against the maximum performance that can be expected (Haag & Borst, 1997; Roddey, Girish, & Miller, 2000). The best model we found, a cascade of two dynamic nonlinearities and a static nonlinearity, is performing close to this maximum.

2. Methods

2.1. Stimuli and measurements

NTSIs were measured with a light detector worn on a headband by a person walking through a natural environment (van Hateren & van der Schaaf, 1996; van Hateren, 1997). The light detector had an acceptance angle of approximately 2 arcmin, and followed the pointing direction of the face (see van Hateren, for further details). Obviously, the NTSIs thus measured are not identical to those that fly photoreceptor cells would normally encounter: many parameters, such as speed and behavior of the organism, average distance to objects, and acceptance angle of the photoreceptors are quite different. It can be argued that several of these parameters cancel each other, and that scale-invariance of the environment will produce NTSIs for different organisms that are not as different as expected (van Hateren), but this argument is not essential for the present study. Here it is only important that the stimulus is sufficiently complex, i.e. that it contains variations in intensity and contrast with the right (natural) mixture of predictability and unpredictability. These variations will then drive the photoreceptor cells into regimes of gain control similar to those in which they function in truly natural circumstances. This makes the stimulus different from laboratory stimuli like sinusoids, flashes, and white noise, which are less complex (lower dimensional, in the sense that they can be generated on the basis of only a few parameters). In van Hateren it was shown that photoreceptors can handle the NTSIs quite well, and that stimulus and response are not linearly related.

The measured NTSIs were played back (at 1200 Hz) on a high-brightness LED, producing light intensities comparable with daylight conditions (van Hateren, 1997). The LED produced a wide-field stimulus of approximately 15° diameter; we found no change in responses for smaller diameters as long as the stimulus covered the receptive field (approximately 2° in diameter) of the photoreceptor cell from which the membrane potential was measured. The LED was driven by a D/A-convertor of a computer connected to a voltage-to-current convertor. The output of the LED was measured with a photodetector with a linear response characteristic; this output was very similar to the original time series, showing only minor nonlinearities in the LED and LED driver. For the analysis in this article, the measured LED output was used as the NTSI, as this was the actual stimulus given to the photoreceptor. The membrane potential of the photoreceptor cell was recorded by using an intracellular microelectrode, it was sampled at 1200 Hz, and stored for off-line analysis. The results presented here are based on 16 measurements in seven photoreceptor cells of four blowflies (*Calliphora vicina*). For further details see van Hateren (1997).

2.2. Model evaluation

Fig. 1 summarizes the methods used for model evaluation, and their relationship with the methods from which they derive. Fig. 1A shows a scheme of the reconstruction method for obtaining the mutual information between stimulus and response of a neuron, as

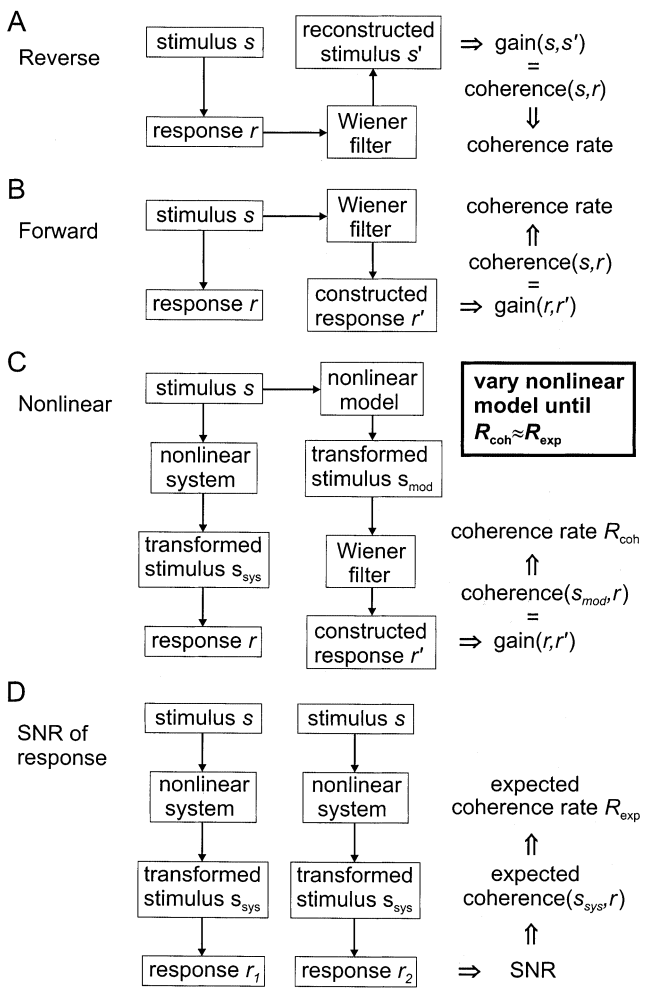

Fig. 1. Coherence rates as a tool for evaluating nonlinear models of a physiological system: (A) In the standard reconstruction method, the stimulus is reconstructed based on the response of the system, using a linear filter; the reconstructed stimulus is compared with the actual stimulus, from which a coherence and a coherence rate (an upper bound on the information rate) are obtained; (B) a forward version of A gives the same coherence and the same coherence rate. Here the response is constructed from the stimulus, again through a linear filter, and compared to the measured response; (C) if the system is nonlinear, the linear method of B can not be applied directly; by making a model that incorporates all nonlinearities of the system, the response can be linearly constructed from the output of the model (yielding r'), and subsequently be compared with the measured response (with r the response to a single presentation of the stimulus s); (D) independent method of obtaining the coherence rate of the system: from a series of responses to the same stimulus, a signal-to-noise ratio is obtained, yielding an expected coherence rate. The better the nonlinear model at C, the closer its coherence rate will be to the coherence rate obtained using the method of D.

pioneered by Bialek, Rieke, Järvilehto, Kouvalainen, Juusola and Weckström (1991; see also Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997), here shown in the frequency domain representation by Theunissen, Roddey, Stufflebeam, Clague, and Miller (1996). Below we will follow the formulation of Haag and Borst (1998). The main idea is to use a Wiener filter to reconstruct the stimulus from its neuronal response. A Wiener filter optimally estimates one signal from another signal with the following assumptions (Papoulis, 1977): the signals are stationary, the estimate is linear, noise is additive, and optimality means here a minimal least-squares error. Then the filter H_{rev} that optimally produces the reconstruction s' of the stimulus, given the response r, becomes (in the frequency domain, all symbols are functions of frequency):

$$H_{\mathrm{rev}} = \frac{\langle sr^* \rangle}{\langle rr^* \rangle}, \tag{1}$$

where the brackets denote ensemble averaging over the spectra s of different time stretches of the stimulus and the spectra r of the corresponding response stretches, and * denotes the complex conjugate. The numerator is the average cross-spectrum of stimulus and response, the denominator the average power spectrum of the response.

Note that Eq. (1) includes ensemble averages of the spectra of time stretches obtained by segmenting the signals. For the analysis in this article segments of 4096 samples (3.41 s) were used; different segment lengths produced very similar results. An alternative to the calculation in Eq. (1) is to compute the various spectra for the entire length of the signals, and then average adjacent frequency components in the spectra (see Bendat & Piersol, 2000, p. 437). We found that this produces virtually identical results to the ones using ensemble averaging. The fact that the results do not depend significantly on segment length suggests that signal and noise can be considered as reasonably stationary.

With Eq. (1) the optimal reconstruction s' from the response r is

$$s' = \frac{\langle sr^* \rangle}{\langle rr^* \rangle} r. \tag{2}$$

The gain $g_{s's}$ relating s' and s is defined as

$$g_{s's} = \frac{\langle s's^* \rangle}{\langle ss^* \rangle} = \frac{\langle \langle sr^* \rangle / \langle rr^* \rangle rs^* \rangle}{\langle ss^* \rangle} = \frac{\langle sr^* \rangle}{\langle rr^* \rangle} \frac{\langle rs^* \rangle}{\langle ss^* \rangle}. \tag{3}$$

The right-hand term of this equation is also known as the coherence γ^2 of s and r (Bendat & Piersol, 2000), therefore $g_{s's} = \gamma^2$.

From the symmetry of the equations, in particular Eq. (3), it is clear that a forward formulation of the problem (Fig. 1B) leads to equivalent equations: the filter H_{fwd} that optimally produces a construction of the response r', given the stimulus s, is

$$H_{\mathrm{fwd}} = \frac{\langle rs^* \rangle}{\langle ss^* \rangle}, \tag{4}$$

with

$$r' = \frac{\langle rs^* \rangle}{\langle ss^* \rangle} s, \tag{5}$$

and

$$g_{r'r} = \frac{\langle r'r^* \rangle}{\langle rr^* \rangle} = \frac{\langle rs^* \rangle}{\langle ss^* \rangle} \frac{\langle sr^* \rangle}{\langle rr^* \rangle} = g_{s's} = \gamma^2. \tag{6}$$

The simplest interpretation of γ^2 is that it is the concatenation (in any order) of the forward and reverse Wiener filters. If the system is linear and noise-free, γ^2 would be one for all frequencies (original signal perfectly recovered after subsequent forward and reverse Wiener filtering). If $\gamma^2 < 1$ then this is due either to noise or other unaccounted signal sources, or to nonlinearities (Bendat & Piersol, 2000). Note that Eq. (6) shows that the coherence functions that follow from the methods of Fig. 1A and B are identical. For a thoughtful discussion of the relative merits of the forward and reverse techniques, see Theunissen et al. (1996).

Fig. 1C shows how we propose to extend the method of Fig. 1B to nonlinear systems. The assumption here is that the stimulus is transformed first by a nonlinear system (yielding s_{sys}) or a nonlinear model (yielding s_{mod}). Both are assumed to be essentially noise-free, i.e. to have an effective noise level small compared to the noise that is added in the final step leading to the response r. This final step is assumed to be linear, and can thus be treated with a Wiener filter as above. The gain $g_{r'r}$ is in this case the coherence between s_{mod} and r. Note that the measured response, r, is the response to a single stimulus presentation, not an average response (Section 2.4).

The final method, shown in Fig. 1D, provides an independent way to obtain a coherence function, and is used in this article as a benchmark for evaluating the models used in the method of Fig. 1C (Section 2.4). Again all nonlinearities are assumed to be included in the nonlinear system, and the final step is taken as linear with transfer function g and assumed additive noise, n, yielding

$$r = g s_{\mathrm{sys}} + n = s_r + n. \tag{7}$$

The coherence between s_{sys} and r is then (again following Haag & Borst, 1998, we will call this the expected coherence, γ^2_{exp}).

$$\gamma^2_{\mathrm{exp}} = \frac{\langle s_{\mathrm{sys}}(g^* s^*_{\mathrm{sys}} + n^*) \rangle \langle (g s_{\mathrm{sys}} + n) s^*_{\mathrm{sys}} \rangle}{\langle s_{\mathrm{sys}} s^*_{\mathrm{sys}} \rangle \langle (g s_{\mathrm{sys}} + n)(g^* s^*_{\mathrm{sys}} + n^*) \rangle} \tag{8}$$

$$= \frac{g^* \langle s_{sys} s^*_{sys} \rangle g \langle s_{sys} s^*_{sys} \rangle}{\langle s_{sys} s^*_{sys} \rangle (\langle s_r s^*_r \rangle + \langle nn^* \rangle)}$$

$$= \frac{\langle s_r s^*_r \rangle}{\langle s_r s^*_r \rangle + \langle nn^* \rangle} = \frac{\mathrm{SNR}}{\mathrm{SNR} + 1},$$

assuming uncorrelated signal and noise, thus $\langle s_{\mathrm{sys}} n^* \rangle = 0$, and with the signal-to-noise ratio, SNR, defined as the ratio of signal and noise power

$$\mathrm{SNR} = \frac{\langle s_r s^*_r \rangle}{\langle nn^* \rangle}. \tag{9}$$

From Eq. (8) it follows that

$$\mathrm{SNR} = \frac{\gamma^2_{\mathrm{exp}}}{1 - \gamma^2_{\mathrm{exp}}}. \tag{10}$$

The expected coherence function $\gamma^2_{\mathrm{exp}}$is determined by the method of Fig. 1D, by first estimating the SNR from a repetition of the stimulus and subsequently applying Eq. (8). The SNR was estimated in the following way. Suppose that the stimulus is repeated m times, yielding photoreceptor responses $\rho_i(t)$ $(i = 1, \ldots, m)$; note that the response is given here as a function of the time t. From these measured responses we obtain the average response

$$\bar{\rho}(t) = \frac{1}{m} \sum_{i=1}^{m} \rho_i(t), \tag{11}$$

and the deviations $\delta_i(t)$ of the responses around the average

$$\delta_i(t) = \rho_i(t) - \bar{\rho}(t). \tag{12}$$

Now we can compute $\bar{r}(\omega)$, the Fourier transform of $\bar{\rho}(t)$, and $d_i(\omega)$, the Fourier transform of $\delta_i(t)$; $\bar{r}$ and d_i are functions of the frequency ω. From this the raw signal and noise power spectra can be computed:

$$S_{\mathrm{raw}} = \langle \bar{r}\, \bar{r}^* \rangle, \tag{13}$$

$$N_{\mathrm{raw}} = \left\langle \frac{1}{m} \sum_{i=1}^{m} d_i d_i^* \right\rangle, \tag{14}$$

where, as before, the brackets denote averages over time segments. Both S_{raw} and N_{raw} are biased estimators, however. The quantity S_{raw} is an overestimate of the power spectrum $S = \langle s_r s^*_r \rangle$ of the actual signal s_r, as used in Eq. (9). A simple example illustrating this problem is when $S = 0$ (i.e. when the photoreceptor output contains only noise), because for finite m the individual responses (consisting only of noise here) are not likely to cancel exactly in the average. Hence $S_{\mathrm{raw}} > 0$, even if $S = 0$. Because part of the noise power is thus mistaken for signal power, the estimated noise power N_{raw} will be an underestimate of the true noise power $N = \langle nn^* \rangle$ as used in Eq. (9). Therefore, estimating the SNR as $S_{\mathrm{raw}}/N_{\mathrm{raw}}$ (as is often done) leads to an overestimate of the actual SNR. Here we correct for this bias as follows. Analogous to Eq. (7), we write the observed responses as a sum

$$\rho_i(t) = \sigma_r(t) + v_i(t) \tag{15}$$

of a noise-free signal σ_r and a noise v_i, assuming σ_r and v_i to be statistically independent. Assuming also that the noises v_i in the repetitions of the experiment are statistically independent, the computed quantity S_{raw} has a mathematical expectation

$$\hat{S}_{\mathrm{raw}} = S + \frac{1}{m} N. \tag{16}$$

From Eq. (11) and Eq. (15) we find

$$\bar{\rho}(t) = \sigma_r(t) + \frac{1}{m} \sum_{i=1}^{m} v_i(t), \tag{17}$$

and consequently with Eqs. (12), (15) and (17)

$$\delta_i(t) = (\sigma_r + v_i) - \bar{\rho} = v_i - \frac{1}{m} \sum_{j=1}^{m} v_j$$

$$= (1 - \frac{1}{m}) v_i - \frac{1}{m} \sum_{\substack{j=1 \\ j \neq i}}^{m} v_j. \tag{18}$$

Because $\mathrm{d}_i(\omega)$ is the Fourier transform of $\delta_i(t)$, we find with Eq. (14) that the computed quantity N_{raw} has a mathematical expectation

$$\hat{N}_{\mathrm{raw}} = \left(1 - \frac{1}{m}\right)^2 N + \frac{m-1}{m^2} N = \frac{m-1}{m} N, \tag{19}$$

again assuming that the noises v_i are statistically independent. From Eq. (16) and Eq. (19) it follows that unbiased estimates $\hat{N}$ and $\hat{S}$ of N and S can be obtained as

$$\hat{N} = \frac{m}{m-1} N_{\mathrm{raw}} \tag{20}$$

$$\hat{S} = S_{\mathrm{raw}} - \frac{1}{m} \hat{N} = S_{\mathrm{raw}} - \frac{1}{m-1} N_{\mathrm{raw}}. \tag{21}$$

Finally, from $\hat{S}$ and $\hat{N}$ an estimate of the SNR is obtained as

$$\mathrm{SNR} = \frac{\hat{S}}{\hat{N}} = \frac{m-1}{m} \frac{S_{\mathrm{raw}}}{N_{\mathrm{raw}}} - \frac{1}{m}. \tag{22}$$

For very large m, Eq. (22) reduces to the raw estimate $S_{\mathrm{raw}}/N_{\mathrm{raw}}$, but for small m it gives a better estimate of the actual SNR.

2.3. The coherence rate

If signal and noise have Gaussian statistics, and are independent of each other, Shannon's equation (Shannon, 1948) gives the information rate R in the channel as

$$R = \int_0^\infty \log_2(1 + \mathrm{SNR})\, \mathrm{d}f, \tag{23}$$

with f the frequency. With Eq. (10) this gives an information rate

$$R = -\int_0^\infty \log_2(1 - \gamma^2_{\mathrm{exp}})\, \mathrm{d}f. \tag{24}$$

This quantifies, with a single number, how close the coherence function is to one over the entire frequency range. The contribution of frequencies where γ^2_{exp} is close to one is very large, and it is very small for frequencies where γ^2_{exp} is close to 0. It is a useful measure even when the conditions for Shannon's Equation (23) are not met. It then simply summarizes the coherence function in a way that is more meaningful than, for instance, the average of the coherence function over a particular frequency band. In the remainder of this article, this single number will be called the coherence rate R_{coh}; it is given in bit/s as it can be considered as an approximation of the information rate. R_{coh} is defined for any coherence γ^2 as

$$R_{\mathrm{coh}} = -\int_0^\infty \log_2(1 - \gamma^2)\, \mathrm{d}f. \tag{25}$$

Using this definition and the term 'coherence rate' aims to avoid confusion with the true information rate of the system.

The method of Fig. 1D yields the expected coherence γ^2_{exp}. Using this result for γ^2 in Eq. (25), we obtain the expected coherence rate, R_{exp}. It is in fact the expectation of the coherence between single responses r_i and a noise-free system response s_{sys}, where the latter is determined by averaging many responses. The result of this averaging can be considered as the response that the best possible (non-linear) model should give. Any model that deviates from this best possible one will show larger deviations from the measured responses, and thus a coherence rate smaller than the coherence rate, R_{exp}, due to the method of Fig. 1D. This conclusion depends on whether the averaging is not biased by experimental artefacts such as drift in the measurements during the course of the experiment. This appears not to be a significant problem in the present experiments, as we found that the cross-correlation between stimulus and noise (as determined according to the method of Fig. 1D) is negligible. We also did not find a systematic change in noise level as a function of response level. Therefore, R_{exp} can be considered as a target for the coherence rate, an upper bound against which coherence rates obtained through the method of Fig. 1C can be gauged. A variant of this method was recently used by Roddey, Girish, and Miller (2000) for investigating the performance of a linear model for receptor cells in the cricket cercal system.

2.4. General method

The general method followed in this article is then as follows: for a given model, parameters were fitted to maximize the coherence rate R_{coh} for that model (Fig. 1C). This maximization was performed using a simplex algorithm (Press, Teukolsky, Vetterling, & Flannery, 1992). Coherence rates were calculated with Eq. (25) by

integrating up to 200 Hz; this frequency range includes virtually all signal power for blowfly photoreceptor cells. For the results presented below, the coherence functions and responses were calculated for the same full stretch of 300 s data as was used for fitting the parameters of each model. To check against overfitting, we performed control computations where coherence rates were calculated for a different part of the time series than what was used for the fits, and found virtually identical results. Different models were investigated, each time maximizing R_{coh}, in order to find a model with R_{coh} as close as possible to R_{exp} (Fig. 1D). The models presented below are in fact the most instructive or interesting results of this search through model space.

For evaluating the models of Fig. 1C we chose to use the response r to a single stimulus presentation for calculating the coherence function with the model outcome. An alternative would have been to use the average response to a large number of stimulus repeats. We did not use this alternative because it would yield much higher coherence rates (because the noise would be averaged out), which could not be compared with the results of the method of Fig. 1D. Thus we would loose our benchmark. A potential advantage of using averages rather that single responses is that it avoids that the models would try to fit the noise in the responses as well as the signal. The control computations mentioned above already suggest that this effect does not occur here to any significant degree. It is also not to be expected because the stimulus is long (typically 300 s). Even if the model would manage to fit noise in one particular segment of the stimulus, this is likely to be punished by a decreased quality of fit in another segment (where the stimulus and reponses might be roughly similar, but the noise would most likely be different). In this respect a long stimulus is similarly effective in avoiding fitting noise as a much shorter stimulus repeated many times.

3. Results

Below we will first present results from models that contain no or only a static (memory-less) nonlinearity, subsequently from models with a dynamic nonlinearity, and finally from combinations of these. The longer-term behavior is discussed next, and for the purpose of simulations a complete model, including a parametrization of the Wiener filter, is presented.

3.1. Models with a static nonlinearity

The upper two rows of Fig. 2 show the stretch of 300 s of the NTSI used for most of the experiments in this article. The graphs show the intensity with a linear and logarithmic ordinate, with frequencies higher than 2.5 Hz removed for the purpose of presentation; all analysis in this article, however, is performed using the full sampling rate. The amplitude spectrum of an NTSI behaves approximately as $1/f_t^{1/2}$ (with f_t temporal frequency) over an appreciable frequency range (van Hateren, 1997). Thus low temporal frequencies dominate the stimulus, although amplitudes at high frequencies are still appreciable because the fall-off with frequency is relatively slow. The distribution of intensities is quite skew, with values at high intensities sparser than at low intensities (Fig. 2, upper row); taking the logarithm of the intensity leads to a more symmetrical distribution (Fig. 2, second row). Further details on the statistics can be found in van Hateren (1997).

The model results below are based on 16 measurements of the response of photoreceptors to the NTSI shown in Fig. 2 (obtained from recordings of seven photoreceptor cells in four flies). The responses shown in Figs. 2 and 4 are from a typical cell, which had a coherence rate and parameter values for model M_{DWN} close to the average of the entire set of measurements.

The results of three simple models are shown in the third row and below of Fig. 2. The first model, M_{lin}, contains no nonlinearities, so it just consists of the (linear) Wiener filter. The thin line (red) shows the response of the photoreceptor cell, and the thick broken line (blue) the optimal Wiener prediction of the response. Both the total stretch of 300 s (limited to 2.5 Hz) and three shorter sections of 500 ms (at full resolution) are shown. The shorter sections are taken from positions denoted by the vertical bars on the abscissa of the 300 s graph; the bar widths cover approximately the temporal extent of these sections. As is clear from the graphs, the linear model is not performing very well: apart from large DC shifts, also the amplitude of fast modulations is not well predicted. This is also clear from the coherence function shown to the right (obtained through Eq. (6)). A maximum coherence of 0.8 is in fact not particularly high: for a linear system it would correspond, via Eq. (10), to an $SNR = 4$.

The second model shown in Fig. 2, M_{log}, is doing much better. It consists of a static nonlinearity, a logarithm, followed by a Wiener filter. The coherence function is now approximately 0.95 at low frequencies, corresponding to an $SNR = 19$. A logarithm is a very simple implementation of Weber's law: it gives equal responses to stimuli of equal contrast, i.e. stimuli which scale in proportion to the local (time-) average of the intensity. The logarithm is also closely related to the dynamic gain control module called 'Weber' below (M_W).

The final model of Fig. 2, M_{sqrt}, consists of a square-root nonlinearity followed by a Wiener filter. Although it performs somewhat worse than M_{log}, this is mainly due to discrepancies at low frequencies (as can be seen

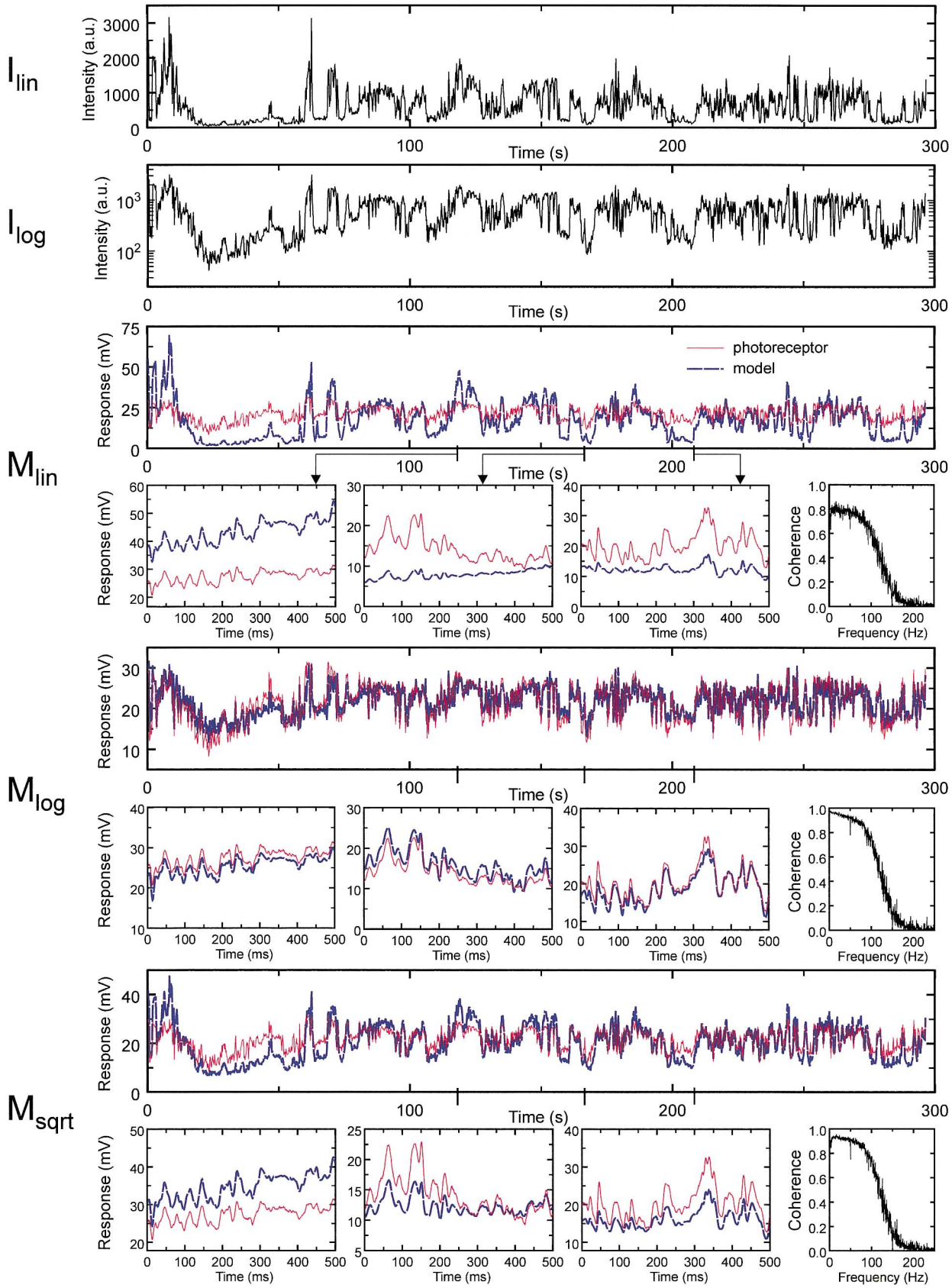

Fig. 2. Upper two rows: 5 min of a Natural Time Series of Intensities (NTSI), drawn with the ordinate linear and logarithmic, respectively. M_{lin}: linear model, where the fat broken trace (blue) shows the optimal Wiener prediction of the measured photoreceptor response (membrane potential, thin red trace). Three sections (bars with arrows) are shown at a higher time resolution. To the right of these graphs the coherence (of model output and measured response) is shown for the 5 min trace. M_{log}: as M_{lin}, after a logarithmic transform of the NTSI. M_{sqrt}: as M_{lin}, after taking the square root of the NTSI.

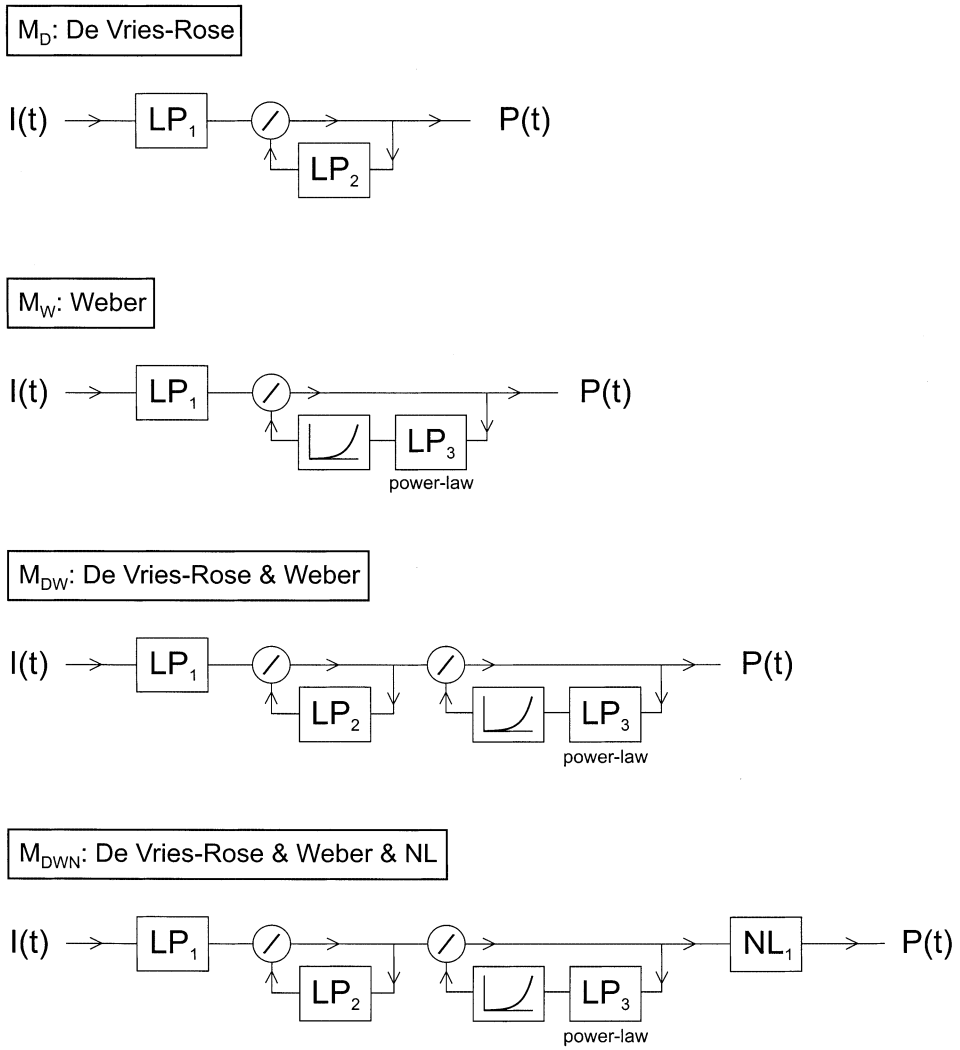

Fig. 3. Models with dynamic nonlinearities. Fixed parameters: M_D: LP_1: $n=3$ (3rd-order low-pass filter), LP_2: $n=1$; M_W: LP_1: $n=3$, LP_3: power $=-0.5$, nonlinearity: $\exp(k_2 \cdot \text{input})$; M_{DW}: LP_1: $n=3$, LP_2: $n=1$, LP_3: power $=-0.5$, nonlinearity: $\exp(k_2 \cdot \text{input})$; M_{DWN}: LP_1: $n=3$, LP_2: $n=1$, LP_3: power $=-0.5$, nonlinearity: $k_1\exp(k_2 \cdot \text{input})$, NL_1: output $=$ input$/(1+$input$)$. See text for further explanation of the models.

in the 500 ms sections). The coherence at higher frequencies is similar to that of M_{log}, which means that the coherence rate Eq. (25) is not much lower for M_{sqrt} than for M_{log} (Section 3.3). The square-root nonlinearity is closely related to the dynamic gain module called 'DeVries-Rose' below (M_D).

3.2. Models with a dynamic nonlinearity and cascaded models

Fig. 3 shows schemes of several models that contain dynamic, rather than static nonlinearities. These models are inspired by a recent model for light adaptation in the human visual system (Snippe, Poot, & van Hateren, 2000). The upper two rows show models with only a single control loop. The DeVries–Rose model, M_D, contains a divisive feedback. Its steady-state behavior follows a square root, because for the steady state output $=$ input/output, thus output $=\sqrt{\text{input}}$. A square-root scaling of sensitivity with luminance is commonly known as the DeVries-Rose law, hence the model name. Note, however, that this behavior of the model is not caused by photon noise (a common cause of square-root behavior), but is just a property of the feedback structure. Because of the low-pass filter LP_2, the model produces overshoots at increment steps of the intensity, and undershoots at decrement steps. Fig. 4, upper two rows, shows the performance of this model. Not surprisingly, it performs similarly to M_{sqrt}, although slightly better (Section 3.3).

The second model of Fig. 3 is the Weber model, M_W, which contains an exponential nonlinearity in the feedback loop. This is similar to Automatic Gain Control (AGC) systems as used for regulating sound or video amplifiers (Ohlson, 1974). In the steady state it gives output $=$ input/exp(output), which yields output $\approx$

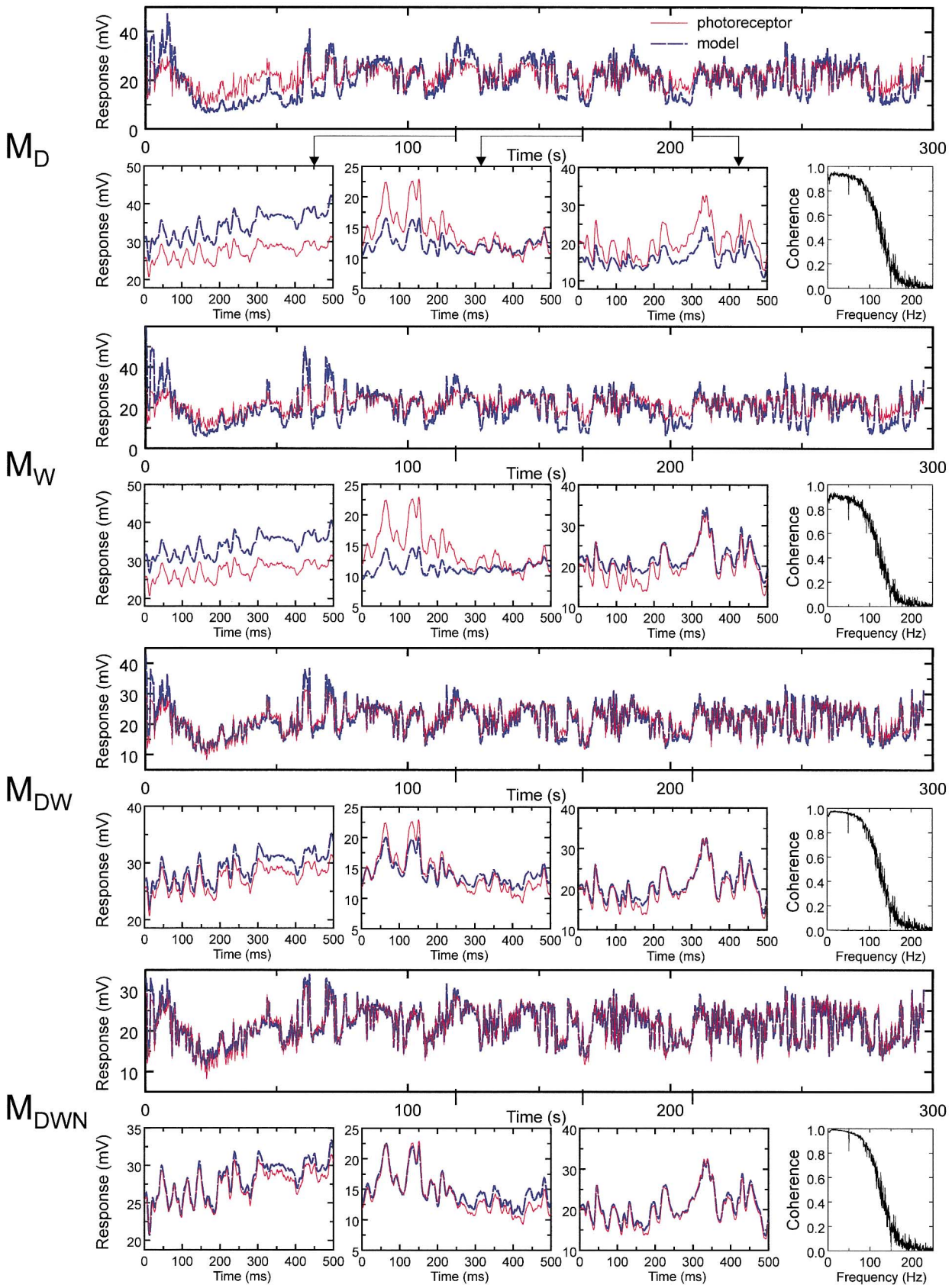

Fig. 4. As Fig. 2, for the four models shown in Fig. 3. Fitted parameters: M_D: LP_1: $\tau = 0.96$ ms, LP_2: $\tau = 8.8$ ms; M_W: LP_1: $\tau = 1.37$ ms, $k_2 = 1.7 \times 10^4$; M_{DW}: LP_1: $\tau = 1.21$ ms, LP_2: $\tau = 6.34$ ms, $k_2 = 2.13 \times 10^3$; M_{DWN}: LP_1: $\tau = 1.76$ ms, LP_2: $\tau = 71.4$ ms, $k_1 = 2.57$, $k_2 = 9.98$.

log(input) if the scaling is chosen such that log(output) ≪ output. Therefore, an AGC loop has, like a pure logarithm, built-in Weber behavior. Although for AGC systems LP_3 is usually taken as a first-order low-pass filter, it was chosen here in a different way. When an intensity step is presented to a blowfly photoreceptor, the response will, after an initial fast transient, continue to drop slowly for quite a long time (seconds to minutes). This long-term adaptation is in fact well matched to the properties of NTSIs: these have power spectra that behave approximately as $1/f_t$ (with f_t temporal frequency) over an appreciable range of frequencies (van Hateren, 1997; van der Schaaf, 1998). This means that NTSIs contain strong slow components, producing very long correlation times. If the underlying, 'local average' light intensity of such an NTSI would have to be estimated, one possibility would be to weight the incoming intensities with a matched filter (Papoulis, 1977), i.e. a filter with an amplitude characteristic $\sim 1/f_t^{1/2}$. Although the filter LP_3 in Fig. 3 in fact operates on the model output rather than on the incoming intensities, this does not affect the basic idea since it is known (van Hateren) that, for NTSIs, the actual photoreceptor output still behaves as approximately $1/f_t^{1/2}$. The filter LP_3 in Fig. 3 is implemented as a superposition of first-order filters covering a range of time scales (Thorson & Biederman-Thorson, 1974). It is designated as 'power-law' because the impulse response of a $1/f_t^{1/2}$-filter has a tail that declines as a power law of time ($\sim 1/t^{1/2}$ for a minimum phase filter, see Kasdin, 1995). The filter is limited here to a total time span of 25 s to ensure it is integrable; changing this length by a factor of two had only small effects on the results. Fig. 4, second pair of rows, shows the performance of M_W. It performs slightly worse than M_D.

The final two models schematized in Fig. 3 are cascades of two dynamic nonlinearities with (M_{DWN}) and without (M_{DW}) a final static nonlinearity. This static nonlinearity was implemented as a Naka-Rushton equation (output = input/(1 + input)), but similar results were obtained with an arctangent function as used before (Snippe, Poot, & van Hateren 2000). As Fig. 4 shows, both models perform quite well, better than any of the other models. The rms-deviation between model and measurement is slightly larger for the lowest response levels than for higher ones: for M_{DWN} the rms-deviation is 1.27 mV for responses < 10 mV (4.5% of the data), 0.92 mV for 10–15 mV (24%), 0.93 mV for 15–20 mV (28%), 0.87 mV for 20–25 mV (24%), 1.05 mV for 25–30 mV (17%), and 0.96 mV for responses > 30mV (2.5%). A similar trend was found for the rms-deviation between model and measurement as a function of stimulus intensity. In the next section we will investigate the relative overall performance of the models, and how well they compare to the maximum performance that can be expected.

3.3. Performance of the models

The coherence functions shown in Figs. 2 and 4 are those according to Eq. (6), following the scheme of Fig. 1C. They show the coherence between the measured response r and the nonlinear model prediction s_{mod}. An independent way to estimate a coherence function is through the scheme of Fig. 1D. The stimulus is here repeated many times, and only the responses are subsequently studied. The signal power spectrum follows from the average response, and the noise power spectrum from the average of the power spectra obtained from the difference between each individual response and the average response. From these spectra the signal-to-noise ratio, SNR, can be obtained through Eq. (15). This subsequently yields, through Eq. (8), an estimate of the expected coherence function, γ^2_{exp}. The expected coherence function can be considered as an upper bound of the coherence function of the system (Section 2.3). As only the repeatability of the response is considered, there are no assumptions involved on a specific nonlinear model linking stimulus to response. Therefore, by comparing this with the coherence functions calculated for the models in Figs. 2 and 4, it is possible to quantify how close the models are to the maximum coherence that can thus be expected. Fig. 5A shows an example of γ^2 for M_{DWN} (fat line) and γ^2_{exp} (thin, grey line) determined for the same photoreceptor cell. As can be seen, they are quite close. An alternative view of the performance is given by $I_{coh} = -\log_2(1-\gamma^2)$, which is perhaps more adequate because the coherence rate is the integral over I_{coh}. From the resulting Fig. 5B it is clear that the remaining discrepancies are mainly in the low-frequency part of I_{coh}, but that they are not very large.

Although Fig. 5A shows that the coherence remains quite high up to frequencies as high as 70–80 Hz, this does not imply that this also holds for the SNR and I_{coh}. The reason is that coherences of similar magnitude, for example 0.99 and 0.95, can be associated with quite different SNRs, in the example 99 and 19, respectively. As Fig. 5B shows, the performance of the photoreceptor to NTSIs already starts to decrease at frequencies of 20 Hz and up, at about the same frequencies where the low-pass filtering of the photoreceptor starts to become apparent (van Hateren, 1997).

Integrating the curves in Fig. 5B over frequency gives two estimates of the coherence rate, firstly the expected rate due to γ^2_{exp}, and secondly the coherence rate following from the γ^2 obtained from M_{DWN}. These numbers were computed for all measurements and models, and the results are shown in Fig. 6. The filled circles show the coherence rate obtained by averaging the rates of the 16 measurements. Error bars show the s.e.m., with most of the bars falling within the boundary of the circles. The coherence rate of the best model (M_{DWN}) is

only slightly smaller than the expected coherence rate (open circle, average and s.e.m. of eight measurements with each typically 15 repeats; measurements were from four photoreceptors cells of three blowflies).

3.4. Longer-term behavior

Most measurements were obtained with an NTSI of 5 min duration. A series of longer recordings of 25 min were obtained as well (seven measurements from six cells in three flies), but intracellular recording time was not long enough to repeat this stimulus often enough to obtain a reliable estimate of γ^2_{exp}. Nevertheless, using the model $\mathrm{M_{DWN}}$, it is possible to study the coherence rate and how it varies over longer times. Fig. 7 shows the 25 min stretch of NTSI used (with the first 5 min identical to the NTSI of Fig. 2; graph limited to frequencies lower than 1 Hz). After about 1050 s the average intensity drops considerably, because for this particular recording a wood was entered after walking through half-open countryside. The third row gives a typical photoreceptor response to this stimulus; this only gives a rough idea of the response, however, because it had to be limited to frequencies lower than 1 Hz for the purpose of presentation. Model $\mathrm{M_{DWN}}$ was applied to each minute of this stimulus-response pair, with different parameter settings fitted to each minute. From this the coherence rate in each minute was calculated, which is shown in the lower graph (mean and s.e.m. of seven measurements). The coherence rate varies somewhat, but most when the average intensity varies. The coherence rate averaged over the 25 min was 460 bit/s.

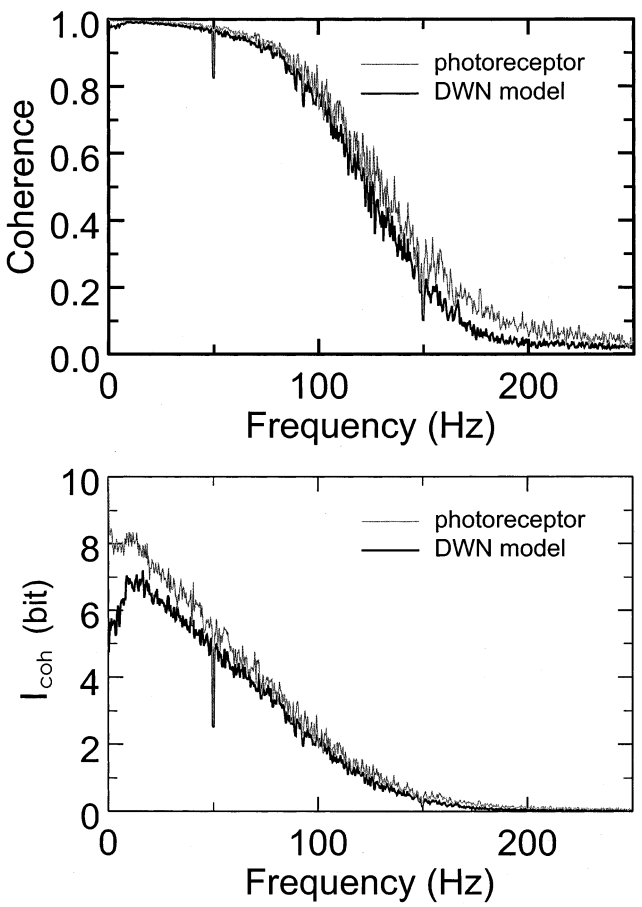

Fig. 5. Comparison of the coherence (upper panel) and $I_{\mathrm{coh}} = -\log_2(1 - \mathrm{coherence})$ (lower panel) obtained with the method of Fig. 1C (dark trace, using $\mathrm{M_{DWN}}$ as in Figs. 3 and 4) and with the method of Fig. 1D (light trace). The latter is labeled 'photoreceptor', as it reflects the signal-to-noise ratio measured directly in the photoreceptor, without any specific assumptions on a model. It gives an upper bound on the coherence that can only be reached by the modelling method of Fig. 1C if both the nonlinear model is adequate and the assumptions underlying the method are correct.

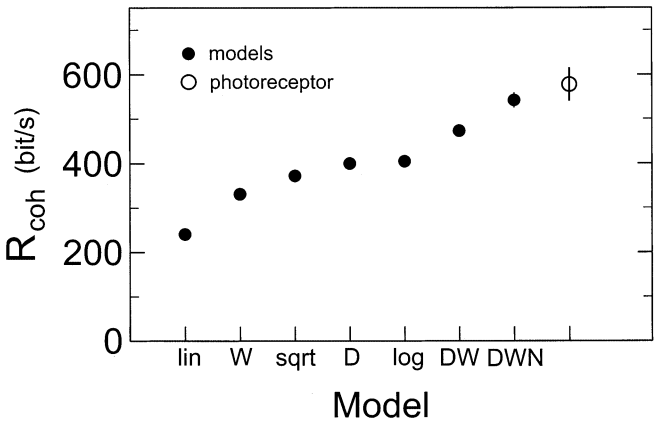

Fig. 6. The coherence rate for a 5 min NTSI of the models evaluated in Figs. 2 and 4, compared to the direct estimation of Fig. 1D (labeled 'photoreceptor'). Average and s.e.m. of 16 measurements (models) and eight measurements ('photoreceptor'). Several error bars are not visible because they fall within the extent of the dots.

The performance of the various models when applied to the entire 25 min (thus now with parameters fixed for the entire period) is shown in Fig. 8. The model $\mathrm{M_{DWN}}$ gives a coherence rate of 389 ± 33 bit/s, appreciably smaller than the 460 bit/s found for separate minutes. Indeed, the parameters obtained from the fits to single minutes vary systematically, in particular with the average light intensity of each minute. Fig. 9 shows an attempt to take this variation into account, by varying the time constant of the first filter, $\mathrm{LP_1}$, depending on the output of filter $\mathrm{LP_3}$. The latter gives a rough, slowly varying estimate of the logarithm of the light intensity. Indeed, this model, M_τ, raises the coherence rate to 415 ± 30 bit/s. Although this may not seem significantly larger than the 389 ± 33 bit/s of $\mathrm{M_{DWN}}$, this is masked by the variation between cells: all seven individual measurements gave an increased coherence rate, with an average increase of $7.4 \pm 2.3\%$. Nevertheless, as this coherence rate still falls short of the 460 bit/s average for single minutes, it is clear that model M_τ only captures part of the longer-term variations of photoreceptor cell functioning.

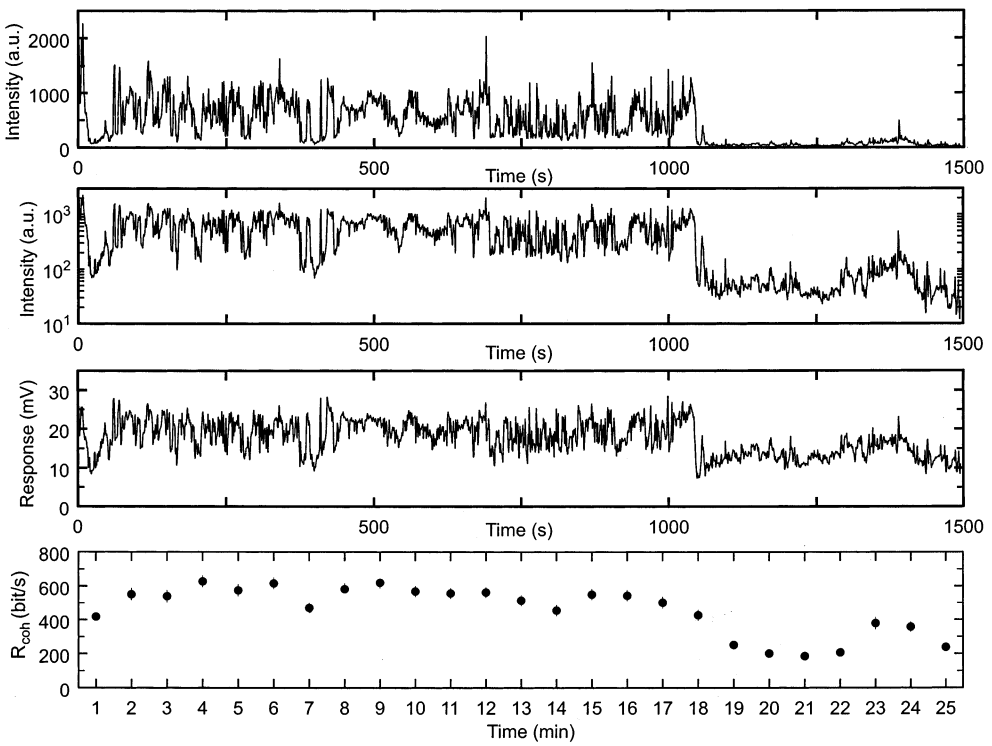

Fig. 7. An NTSI of 25 min (upper two panels), an example of the response of a photoreceptor to this stimulus (third panel, limited to 1 Hz for the purpose of presentation), and the coherence rate of model M_{DWN} evaluated for consecutive minutes (lower panel; average and s.e.m. of seven measurements).

3.5. Models including a parametrization of the Wiener filter

Although the models with static nonlinearities evaluated in Fig. 2 and the dynamic models shown in Fig. 3 take care of the nonlinearities in the system, they have to be followed by a linear filter (the Wiener filter) in order to completely specify the response of the system to a specific stimulus. The Wiener filters that result from the analysis of the various models are found in numerical form. For the purpose of evaluating the models or using them as preprocessing modules in future studies of higher visual processing, it is more convenient to summarize them with a simple function. We will do that here for two of the models, M_{sqrt} and M_{DWN}. The first is the simplest model that performs reasonably well for both the 5 and the 25 min NTSI (Figs. 6 and 8). Model M_{DWN} is the overall top-performer. Fig. 10A shows the Wiener filter for M_{sqrt} (dots) and its descriptive function (line). Fig. 10B shows this for M_{DWN}. Finally, Fig. 10C and D show the response of M_{DWN} to several simple stimuli that have often been used for testing light-adapted photoreceptors: the responses to a very short pulse and to a series of 500 ms intensity steps of various magnitudes, both decrements and increments. In Section 4 these model responses will be further assessed.

4. Discussion

In this article we used an information theoretical method to compare the performance of a range of models describing light adaptation and gain control in blowfly photoreceptor cells. The method makes it possible to gauge the performance of the models in terms of a coherence rate against an upper bound of a coherence rate that was estimated from the repeatability of the response to the same stimulus. The main conclusions are that nonlinearities are needed for an adequate model, and that dynamic nonlinearities perform better

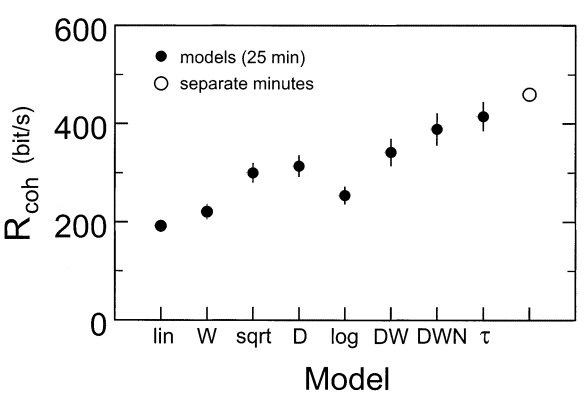

Fig. 8. The coherence rate for a 25 min NTSI of the models evaluated in Figs. 2 and 4, compared to the average rate obtained by fitting M_{DWN} to each minute separately (data points of Fig. 7, lower panel).

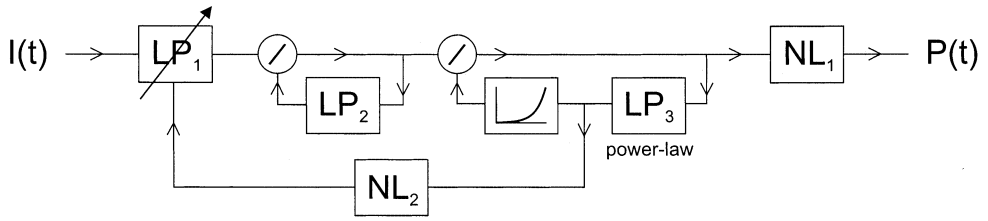

Fig. 9. M_τ, a modification of M_{DWN} with a variable time constant of the first low-pass filter. Fixed parameters: LP_1: $n=3$, LP_2: $n=1$, LP_3: power $=-0.5$, nonlinearity: $k_1\exp(k_2\cdot\text{input})$, NL_1: output $=$ input$/(1+$input$)$, NL_2: $\tau_1=\tau_0/(\text{input})^w$. Fitted parameters for the measurement of Fig. 7, third row: LP_1: τ_1 output of NL_2, with $\tau_0=0.28$ ms, LP_2: $\tau=43.3$ ms, $k_1=8.18$, $k_2=7.18$, $w=1.52$.

than static (memory-less) nonlinearities. Finally, for models describing longer-term behavior additional control loops, adjusting for instance the time constants of the system, appear to be necessary.

The final model, M_{DWN}, consists of a cascade of a fast feedback gain control loop, a slow nonlinear feedback loop, and a static nonlinearity. The first feedback loop, behaving approximately as a square-root device, will significantly reduce the considerable dynamic range of the stimulus (short-term range 10^2–10^3) down to a size (10–30) that is manageable by the physiology of the photoreceptor cell. Larger steps in stimulus intensity, occurring predominantly on much slower time-scales, are handled by the second control loop, containing a power-law low-pass filter matching the time-scale invariance of NTSIs. Finally, any remaining peaks that would drive the photoreceptor cell out of its dynamic range, are handled by the final static nonlinearity (van Hateren & Snippe, 2000). This combined system keeps the response well within the dynamic range of the photoreceptor cell, and produces a coherence rate close to the upper bound estimated from the repeatability of the response.

The number of fitted parameters is 4 for M_{DWN}, with several more implicitly fixed by the choice of filters and functional forms of nonlinearities. Clearly, the coherence rates following from the progression of models in Figs. 2 and 4 increase with the number of parameters. Nevertheless, given the complexity and high dimensionality of the stimulus and response, the number of parameters is very modest in the nonlinear part of the models. Although the final (linear) Wiener filter does have many degrees of freedom, these do not contribute to the coherence estimates. This can be readily shown from Eq. (6): multiplying s (or r) by an arbitrary linear filter F does not affect the coherence, because the effects of F in the numerator and denominator of Eq. (6) cancel. The Wiener filter in the model only serves to minimize the root-mean-square deviations between the model outputs and the observed photoreceptor response.

The M_{DWN} model is similar to the first stage of the model by Snippe, Pool, and van Hateren (2000), developed for describing light adaptation and contrast gain control in the human visual system. The main difference is in the low-pass filter in the second control loop, which was taken as a very slow first-order filter in Snippe et al. (in fact not dynamically active in the experiments described there). Here it is modelled as a power-law low-pass filter, weighting the incoming NTSI simultaneously over a range of time scales. The result is that adaptation works over a range of time-scales, including quite slow ones. This is consistent with the observed dynamics of fly photoreceptor adaptation.

Several of the components of the model have been used before by Lankheet, van Wezel, Prickaerts, and van de Grind (1993) to describe gain control in horizontal cells in cat retina. They found that their results could be well described by a divisive feedback generating a square-root behavior, followed by a static nonlinearity. A difference with the M_{DWN} model evaluated here is that their model does not have the second, nonlinear feedback stage with a power-law low-pass filter. This module is used here to describe slow components in the gain control.

French, Korenberg, Järvilehto, Kouvalainen, Juusola, and Weckström (1993) developed models for predicting the response of blowfly photoreceptor cells to steps of intensity. These models consist of a cascade of static nonlinearities and a linear filter, thus somewhat resembling the models with static nonlinearities discussed in Fig. 2. Contrary to the present study they subtract the DC-term from the membrane potential, and thus only study the range of temporal frequencies contained in intensity steps of 200 ms duration. It remains to be determined how well their class of models can perform with NTSIs, including variations in average light level on quite slow time-scales.

The responses predicted by the M_{DWN} model to simple stimuli (Fig. 10C and D) are reasonably close to those measured before for light-adapted blowfly photoreceptors (French et al., 1993; Juusola, Kouvalainen, Järvilehto, & Weckström, 1994). A discrepancy is that the V-logI curve of the model (DC membrane potential as a function of the logarithm of the light intensity) is too shallow (Laughlin & Hardie, 1978). This is probably a major cause of the dependence of the rms-deviation on the response level (Section 3.2). Nevertheless,

the performance of the model is quite good if one takes into account that it was in fact developed for and tuned to an entire different class of stimuli (NTSIs). It is quite likely that careful adjustment of the form of the various filters and nonlinearities in the model can significantly improve the reponse to stimuli as in Fig. 10, without deteriorating its performance with NTSIs.

Although the expected coherence rate found here (577 ± 30 bit/s) should not be equated with the true information rate (Section 2.3), it is still interesting to compare these results with those obtained by de Ruyter van Steveninck and Laughlin (1996). They used white noise stimuli of much lower effective contrast, producing response amplitudes not exceeding a few millivolt. In that case the assumptions on linearity of the signal transfer and Gaussian properties of the signals are better fullfilled than here, and the information rate they find for this stimulus is roughly 400 bit/s (estimated from their Fig. 2). They subsequently extrapolate this to higher contrasts, yielding 1000 bit/s. Given the assumptions involved in that procedure, and the non-Gaussian statistics of the stimuli used in the present study, it is clear that their results and those reported here are in essence consistent.

4.1. Coherence rates as a tool for model evaluation

The coherence and coherence rate were used here as a tool to evaluate the adequacy of particular models, as compared to an upper bound on the coherence rate obtained by analyzing the repeatability of the response to the same stimulus. This method provides the modeller with a single number that summarizes how one model compares to another. Moreover, it also shows how far a model is from the final goal, a model equally adequate as the system under consideration itself. Although this can also be accomplished by the more traditional approach of calculating the root-mean-square difference between model prediction and measurement, we believe the present method has several advantages, in particular for modelling information processing systems. Deviations between model and measurement are weighted according to how much information the various frequency bands carry about the signal. Furthermore, the resulting coherence rate has a simple interpretation in that it is an estimate of the information transferred by the system, be it that the estimate will be biased in the case of non-Gaussian signals.

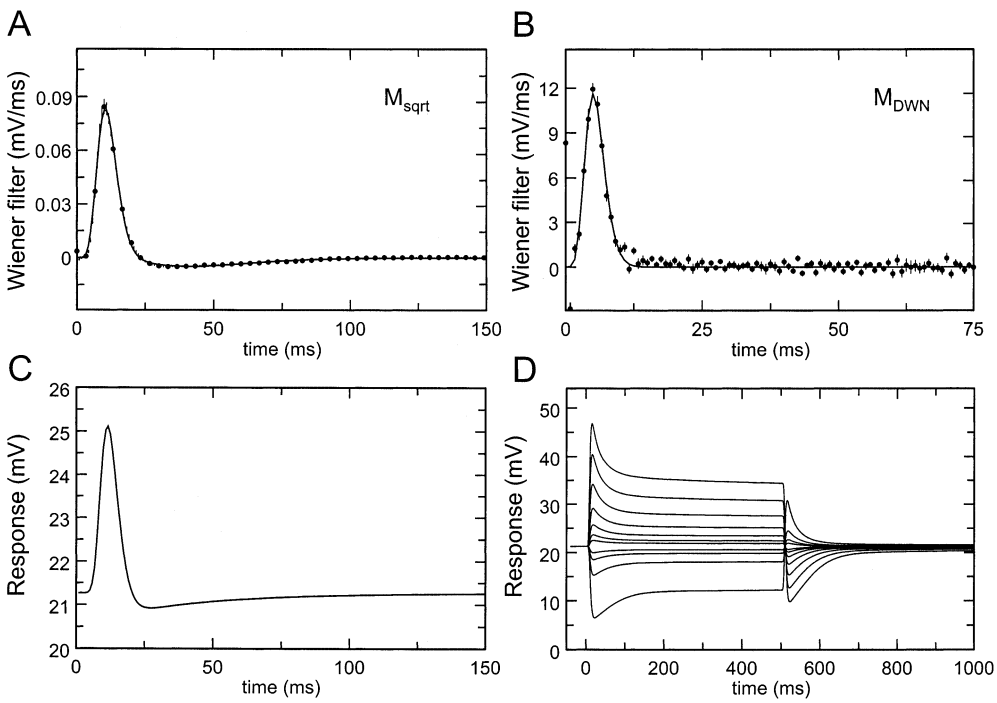

Fig. 10. (A) Wiener filter and its parametric model fit for M_{sqrt}. The dots and error bars show the average Wiener filter and the s.e.m. of 16 measurements, the line shows a fit with the function $A_1(t/\tau_1)^{n_1}\exp(-t/\tau_1) - A_2(t/\tau_2)^{n_2}\exp(-t/\tau_2)$, with $A_1 = 1.85 \times 10^{-7}$ mV/ms, $\tau_1 = 1.133$ ms, $n_1 = 10$, $A_2 = 2.30 \times 10^{-4}$ mV/ms, $\tau_2 = 8.50$ ms, $n_2 = 5$. For the sake of clarity only one in four dots is shown. (B) As A, for model M_{DWN}; the fit is given by $A(t/\tau)^n\exp(-t/\tau)$, with $A = 2.46 \times 10^{-6}$ mV/ms, $\tau = 0.535$ ms, $n = 11$. (C) Pulse response of M_{DWN} (including the parametric Wiener filter) based on the model fit to one of the measured photoreceptor cells. Parameters of the Wiener filter: $A = 3.13 \times 10^{-6}$ mV/ms, $\tau = 0.535$ ms, $n = 11$; further model parameters LP_1: $\tau = 1.69$ ms, LP_2: $\tau = 71.8$ ms, $k_1 = 0.689$, $k_2 = 9.07$. (D) Responses of M_{DWN} to 500 ms steps in light intensity, parameters as in C. Contrast steps: -0.8, -0.4, -0.2, -0.1, 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4.

For the methods of Fig. 1C and D it was assumed that the nonlinear system could be split into a deterministic part followed by an additive noise source. But we know that noise tends to be generated at practically any stage of a real physiological system, thus it remains to be seen how good this assumption is. For a system with distributed noise sources, it depends on the type of nonlinearities how much of the resulting noise can in effect be considered as additive, rather than, for example, multiplicative.

Acknowledgements

We would like to thank Sietse van Netten en Doekele Stavenga for useful comments. Portions of this work were presented at the annual meeting of the Association for Research in Vision and Ophthalmology (van Hateren & Snippe, 2000). The research was supported by the Netherlands Organization for Scientific Research (NWO) through the Research Council for Earth and Lifesciences (ALW).

References

Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Computation*, *2*, 308–320.

Bendat, J. S., & Piersol, A. G. (2000). *Random data: analysis and measurement procedures* (3rd). New York: Wiley-Interscience.

Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R., & Warland, D. (1991). Reading a neural code. *Science*, *252*, 1854–1857.

de Ruyter van Steveninck, R. R., & Laughlin, S. B. (1996). The rate of information transfer at graded-potential synapses. *Nature*, *379*, 642–645.

French, A. S., Korenberg, M. J., Järvilehto, M., Kouvalainen, E., Juusola, M., & Weckström, M. (1993). The dynamic nonlinear behavior of fly photoreceptors evoked by a wide rage of light intensities. *Biophysical Journal*, *65*, 832–839.

Haag, J., & Borst, A. (1997). Encoding of visual motion information and reliability in spiking and graded potential neurons. *Journal of Neuroscience*, *17*, 4809–4819.

Haag, J., & Borst, A. (1998). Active membrane properties and signal encoding in graded potential neurons. *Journal of Neuroscience*, *18*, 7972–7986.

Juusola, M., Kouvalainen, E., Järvilehto, M., & Weckström, M. (1994). Contrast gain, signal-to-noise ratio, and linearity in light-adapted blowfly photoreceptors. *Journal of General Physiology*, *104*, 593–621.

Kasdin, N. J. (1995). Discrete simulation of colored noise and stochastic processes and $1/f^{\alpha}$ power law noise generation. *Proceeding of the IEEE*, *83*, 802–827.

Lankheet, M. J. M., van Wezel, R. J. A., Prickaerts, J. H. H. J., & van de Grind, W. A. (1993). The dynamics of light adaptation in cat horizontal cell responses. *Vision Research*, *33*, 1153–1171.

Laughlin, S. B., & Hardie, R. C. (1978). Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly. *Journal of Comparative Physiology*, *128*, 319–340.

Ohlson, J. E. (1974). Exact dynamics of automatic gain control, IEEE Transactions on Communications COM-22, pp. 72–75.

Papoulis, A. (1977). *Signal analysis*. New York: McGraw-Hill.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in Fortran*. New York: Cambridge University Press.

Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., & Bialek, W. (1997). *Spikes: exploring the neural code*. Cambridge: MIT Press.

Roddey, J. C., Girish, B., & Miller, J. P. (2000). Assessing the performance of neural encoding models in the presence of noise. *Journal of Computational Neuroscience*, *8*, 95–112.

Shannon, D. E. (1948). The mathematical theory of communication. *Bell System Technical Journal*, *27*, 3–91.

Snippe, H. P., Poot, L., & van Hateren, J. H. (2000). A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. *Visual Neuroscience*, *17*, 449–462.

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society London B*, *216*, 427–459.

Theunissen, F., Roddey, J. C., Stufflebeam, S., Clague, H., & Miller, J. P. (1996). Information theoretic analysis of dynamical encoding by four identified primary sensory interneurons in the cricket cercal system. *Journal of Neurophysiololgy*, *75*, 1345–1364.

Thorson, J., & Biederman-Thorson, M. (1974). Distributed relaxation processes in sensory adaptation. *Science*, *183*, 161–172.

van der Schaaf, A. (1998). Natural image statistics and visual processing. Thesis, University of Groningen

van Hateren, J. H. (1992). Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *Journal of Comparative Physiology A*, *171*, 157–170.

van Hateren, J. H. (1997). Processing of natural time series of intensities by the visual system of the blowfly. *Vision Research*, *37*, 3407–3416.

van Hateren, J. H., & Snippe, H. P. (2000). A parametric model for the processing of natural time series of intensities by blowfly photoreceptor cells. *Investigative Ophthalmology & Visual Science*, *41*(Suppl.), 492.

van Hateren, J. H., & van der Schaaf, A. (1996). *Temporal properties of natural scenes. Proceedings of the IS&T/SPIE conference on electronic imaging: science & technology*, vol. 2657 (pp. 139–143). San Jose: SPIE.