Modelling and Prediction of Non-Stationary Optical Turbulence Behaviour

Niek Doelman[a,b] and James Osborn[c]

[a]TNO Technical Sciences, Stieltjesweg 1, Delft, The Netherlands.
[b]Leiden Observatory, Leiden University, Niels Bohrweg 2, Leiden, The Netherlands.
[c]Department of Physics, Durham University, South Road, Durham, United Kingdom.

ABSTRACT

There is a strong need to model the temporal fluctuations in turbulence parameters, for instance for scheduling, simulation and prediction purposes. This paper aims at modelling the dynamic behaviour of the turbulence coherence length r_0, utilising measurement data from the Stereo-SCIDAR instrument installed at the Isaac Newton Telescope at La Palma. Based on an estimate of the power spectral density function, a low order stochastic model to capture the temporal variability of r_0 is proposed. The impact of this type of stochastic model on the prediction of the coherence length behaviour is shown.

Keywords: atmospheric turbulence, coherence length, non-stationary behaviour, stochastic model, forecasting

1. INTRODUCTION

Wavefront phase fluctuations of optical waves propagating through a (weakly) turbulent atmosphere can be regarded as a spatio-temporal stochastic process. Essential descriptive parameters for this stochastic process are for instance: the structure constant of the refraction index fluctuations $C_n^2(h)$, the seeing ϵ_0, the isoplanatic angle θ_0, the coherence length r_0, the coherence time τ_0, the outer scale L_0, the inner scale l_0 and also the wind vector $\vec{v}_w$; see for instance Hardy[1] for the underlying definitions. In the remainder of this paper these parameters will be referred to as *the turbulence parameters*. As an example of the descriptive function of the turbulence parameters, we refer to the well-known von Kármán expression for the 3-dimensional power spectral density of fluctuations in the refractive index:[2]

$$\Phi_n(\kappa) = \frac{0.033C_n^2}{\left(\kappa^2 + \kappa_0^2\right)^{11/6}} \exp\left\{-\frac{\kappa^2}{\kappa_m^2}\right\} \tag{1}$$

in which κ the spatial frequency, $\kappa_0 = 2\pi/L_0$ and $\kappa_m = 5.92/l_0$.

It is well-known from extensive and widespread turbulence profiling campaigns that the turbulence parameters are not static and show time-varying behaviour; see for instance the reference proceedings.[3] An illustration of the temporal variability of the $C_n^2(h)$-profile as measured by the Stereo-SCIDAR instrument is given in Figure 1. Temporal fluctuations of the turbulence parameters imply that the spectrum of index-of-refraction fluctuations (1) - and all properties derived thereof - are also time-varying. This variability of turbulence parameters occurs on various times scales: minutes to hours, diurnally and seasonally.

Having a time-variant spectrum such as for the 3-dimensional fluctuations in refractive index, implies that the resultant stochastic process of wavefront phase fluctuations is non-stationary. Strictly speaking, the spectral representation of (1) is only valid if the turbulence parameters C_n^2, L_0 and l_0 would be time-invariant. We might however, interpret (1) as the instantaneous spectrum in the case of piece-wise constant turbulence parameters. Generally, in literature the temporal variability of the turbulence parameters is represented by a probability distribution, which can be constructed from a series of measurements. Such a distribution reveals the mean,

---

Further author information: (Send correspondence to Niek Doelman)
Niek Doelman: E-mail: niek.doelman@tno.nl, Telephone: 31 888 668 006
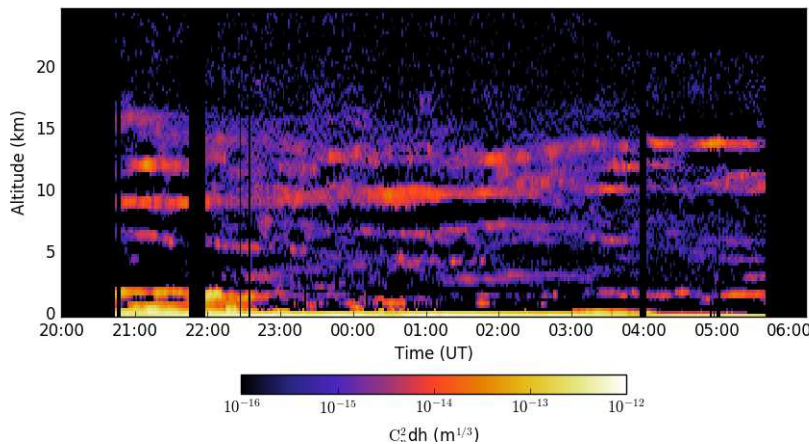
Figure 1. Example C_n^2 profile from the 2.5m Isaac Newton Telescope, La Palma, as measured by the Stereo-SCIDAR instrument on the 30th of June, 2015.

variance and potential higher order moments of the statistical behaviour of the underlying parameter, but it does not show the correlation or even the relation between subsequent time instants. A dynamic model for the temporal evolution of the given variable would contain that information however.

This leads us to the main research question of this paper:

Which time series model can adequately describe the dynamic behaviour of turbulence parameters?

Modelling and also forecasting of the temporal behaviour of the turbulence parameters is essential for the following reasons:

(1) the efficient scheduling of scientific programs and use of astronomical instruments,

(2) the tuning of AO systems and science instruments,

(3) the simulation of astronomical instrument performance under time-varying turbulence.

Furthermore, turbulence modeling and prediction can be used to:

(4) improve the adaptive tracking properties of high-performance controllers - such as predictive or Linear Quadratic Gaussian (LQG)-type controllers - as they need to be tuned towards the specific turbulence parameters; see for instance the on-sky AO control test performed at the McMath-Pierce solar telescope.[4]

1.1 Approach with Stereo-SCIDAR data

The research approach followed here is analysing data from a turbulence profiling instrument based on the SCIntillation Detection And Ranging (SCIDAR) technique. Emphasis is put on the dynamic properties of height-integrated turbulence parameters over relatively short horizons, i.e. up to several hours. Diurnal effects and seasonal effects are initially left out of this analysis.

The Stereo-SCIDAR instrument from Durham University[5,6] uses a stereoscopic system with the scintillation pattern from each star of a double-star target incident on a separate detector. The distance to a turbulent layer is estimated by measuring the offset of a covariance peak in the cross-correlation function of the scintillation pattern from the two stars. The strength of the turbulence is proportional to the magnitude of the covariance. The velocity of the turbulent layers can be estimated optically by tracking the covariance peaks through the spatio-temporal covariance function. By separating the scintillation patterns onto separate detectors instead of overlapping them on a single camera (as with traditional SCIDAR instruments) the noise in the profile

estimation is reduced. This is because, in conventional SCIDAR instruments, the intensity speckles loose contrast in the overlapping patterns, which reduces the visibility of the covariance peaks.

The Stereo-SCIDAR instrument has been installed on the 2.54m Isaac Newton Telescope (INT), La Palma in the past years. It has produced profile data of the optical turbulence strength profile, $C_n^2(h)$ and the optical turbulence velocity profile, $V(h)$. For the purpose of this study data from 20 measurement nights have been used, from July 2014, October 2014 and June/July 2015.

2. ESTIMATION OF THE POWER SPECTRAL DENSITY

To explore the dynamic behaviour of turbulence parameters, we commence with the 'altitude-integrated' coherence length r_0, as defined by

$$r_0 = \left(0.423 \left(\frac{2\pi}{\lambda} \right)^2 \cos(\gamma)^{-1} \int C_n^2(h) dh \right)^{-\frac{3}{5}}, \tag{2}$$

in which γ is the zenith angle and h is the altitude variable. Using the optical turbulence strength profile from the Stereo-SCIDAR instrument the coherence length can be readily determined by integrating over the altitude h.

Given data from 20 observation nights at the 2.54m INT, the distribution of measured r_0 values is as given in Figure 2. The distribution shows resemblance with the log-normal distribution which is a common probability density model for the seeing.[7] Taking the logarithm of the coherence length indeed leads to an approximate normal distribution; see Figure 3. For the given measurements the mean and variance of the normal distribution fit on the $\log(r_0)$ data are respectively: -2.22 and 0.06.
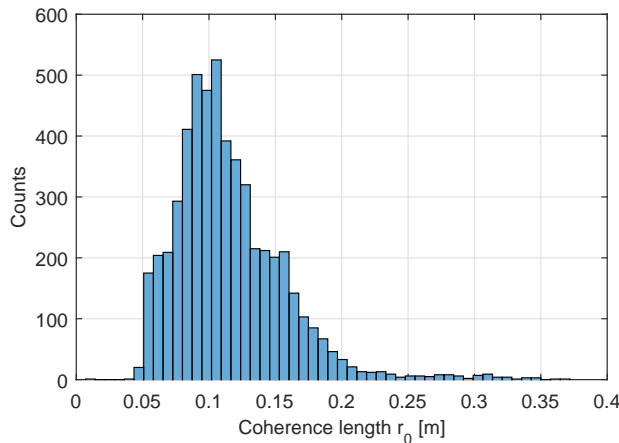

Figure 2. Distribution of coherence length (r_0) values for the 20 measurement nights at INT telescope, La Palma.
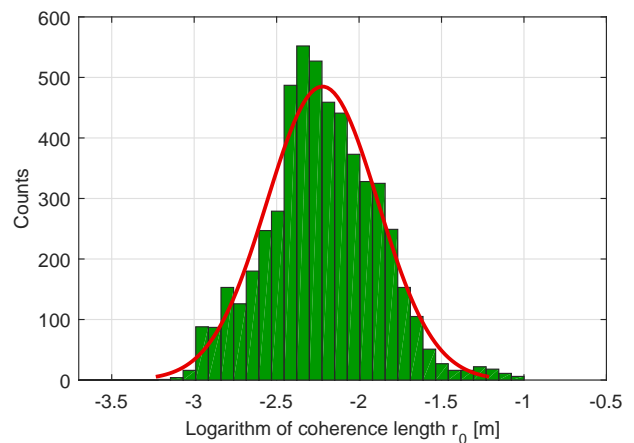

Figure 3. Distribution of the logarithm of the coherence length together with normal distribution fit

The first step is to estimate the power spectral density (PSD) function of the coherence length, based on data from 20 measurement nights at La Palma. Having an accurate estimate of the PSD enables the estimation of a dynamic model for r_0 and subsequently the simulation and prediction of its temporal behaviour. For the particular case of the La Palma data, estimation of the PSD of r_0 is not straightforward. Several factors complicate the estimation approach:

– non-uniform sampling; the time between 2 subsequent measurements is not constant. The majority of sampling intervals range from 80s to 120s.

– missing data; for certain nights periods of 1-2 hours occur without any measurement.

– the recorded time series of r_0 are relatively short and have variable length; in the order of 100 to 400 samples per night.

- the time series consists of non-negative numbers only.
- the variance of the r_0 series shows a large variation per night; this may indicate non-stationarity of the underlying process; see Figure 4.
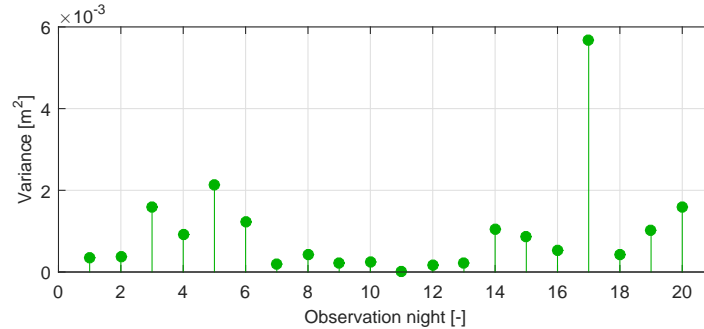

Figure 4. Variance of r_0 values per observation night as measured at INT telescope, La Palma.

2.1 Aliasing Effects

For uniformly sampled time series the valid frequency band for spectral analysis is well-defined by the Nyquist –Shannon sampling theorem. Here, due to the non-uniform sampling and also the missing data intervals it is not straightforward to determine a valid frequency band for the PSD analysis. To assess the potential aliasing effects in this case we consider the spectral window.[8] For an arbitrarily sampled process at N time instants $\{t_i\}_1^N$ the effective spectral window $W(\nu)$ is given by :

$$W(\nu) = \frac{1}{N^2} \left| \sum_{i=1}^{N} \exp(j2\pi\nu t_i) \right|^2 , \tag{3}$$

where ν is the frequency. In the case of uniform sampling $W(\nu)$ is a periodic function, in which the period is two times the Nyquist frequency.

Evaluation of the spectral window for the specific sampling scheme in each observation night and the consequent aliasing effects has led to the results illustrated in Figure 5. From this it can be concluded that at a 'virtual' Nyquist frequency of 1.5 mHz, the aliasing effects for most nights are small (between 1 and 2 dB). Still, the measurement nights 1 and 3 have a significantly higher aliasing effect and these should therefore be removed from the data set. Note, that the effective aliasing above 1.5 mHz grows with the frequency band for each observation night.
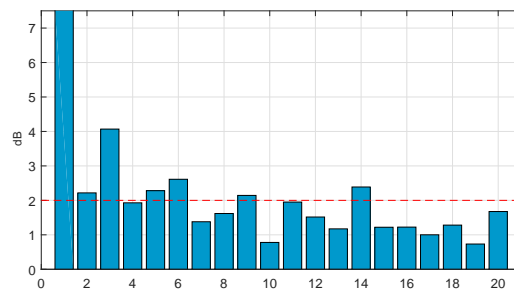

Figure 5. Effective aliasing (in dB's) at 1.5 mHz for each observation night.

2.2 Coherence length PSD estimate

For the r_0 series of the 18 remaining nights the logarithm is taken and the mean is removed. Then, the individual periodogram for each night is estimated and the mean of the periodograms is taken to arrive at an estimate

of the power spectral density function for the temporal fluctuations of $\log(r_0)$. This average PSD is shown in Figure 6. The frequency resolution for the PSD has been determined by the shortest overall measurement span of the 18 observation nights. Note, that the PSD estimate is also shown for frequencies exceeding the selected Nyquist frequency of 1.5 mHz.
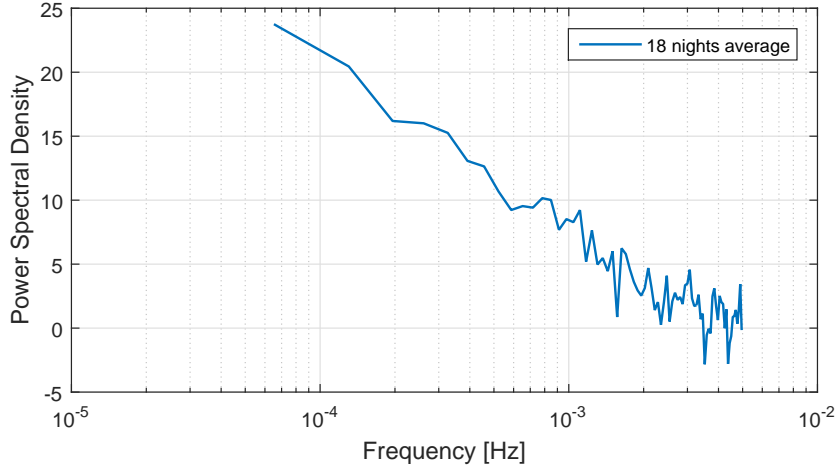

Figure 6. Estimate of power spectral density function based on 18 measurement nights.

3. LOW ORDER STOCHASTIC MODEL

For the analysis of the estimated power spectrum of the coherence length Φ_{r_0} and the matching with a dynamic model, we should focus on the frequency band up to 1.5 mHz. The PSD clearly shows a linear fall-off with the logarithm of the frequency. Such behaviour could be captured by a function $\Phi_{r_0}(\nu) \sim \nu^{-\beta}$. The slope (β) of the PSD function is an essential parameter as it will reveal:

(1) whether the r_0 process has long range or short range dependency,

(2) whether the process is non-stationary and

(3) the horizon over which the coherence length can be accurately predicted.

To match the linear fall-off characteristic in the PSD, a few types of stochastic models might be suitable. An Auto-Regressive-Moving-Average (ARMA) model is a widely used representation of a linear stochastic process. However, as the PSD in Figure 6 shows a fall-off which is not a multiple of 20 dB's/decade, an ARMA model might need a high number of coefficients.

Instead, we propose a low order, fractionally integrated stochastic model of the following form:

$$(1 - q^{-1})^d x_r(k) = \epsilon(k), \tag{4}$$

with d a real-valued - not necessarily integer - parameter, $\epsilon(k)$ a discrete-time white noise process, consisting of independent and identically distributed random variables with zero mean and finite variance σ_ϵ^2, q^{-1} the backward shift operator and $x_r(k)$ the relevant variable as a function of sampling instant k. Note, that in this model representation we do assume uniform sampling. The power spectral density function with model (4) is given by:

$$\Phi_x(\omega) = \frac{\sigma_\epsilon^2}{2\pi} \left| 1 - \exp(-j\omega) \right|^{-2d}, \tag{5}$$

where ω is the (normalised) angular frequency. An alternative expression is: $\Phi_x(\omega) = \frac{\sigma_\epsilon^2}{2\pi} \left| 2 \sin(\frac{\omega}{2}) \right|^{-2d}$.

In the limit case of small ω the PSD function can be approximated by $\Phi_x(\omega) \sim \frac{\sigma_\epsilon^2}{2\pi} |\omega|^{-2d}$ for $\omega \to 0$. Note, that

the PSD has a singularity at $\omega = 0$ for positive d. For $d \in (-\frac{1}{2}, \frac{1}{2})$ the process (4) is stationary and invertible, for $d \geq 1/2$ the process is non-stationary; see for instance Hurvich.[9]

In Figure 7 the spectral match between the r_0 measurements and the model (4) is illustrated for the best fit of d.
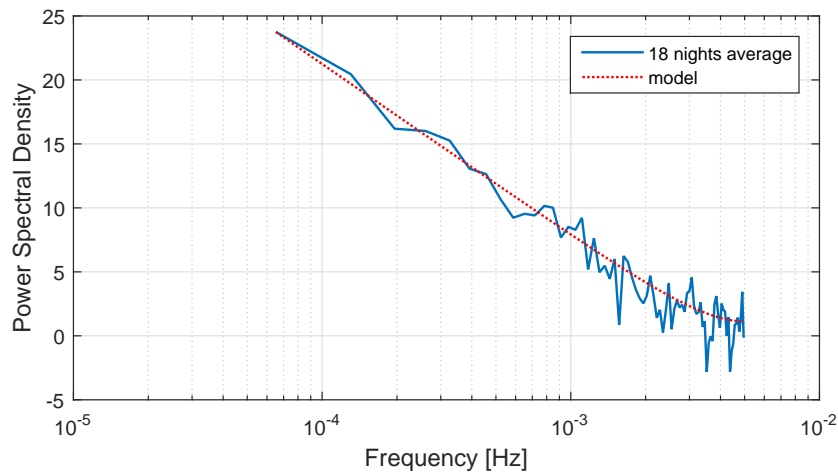

Figure 7. Low order model fit with estimated power spectral density function.

3.1 Estimation of fractional power parameter d

The starting point to estimate d is the average PSD - as shown in Figure 6 - together with the expression for the model PSD (5). By taking the log of the periodogram, the fractional power d can be estimated by a straightforward linear regression approach. This approach belongs to the class of semi-parametric estimators in the frequency domain and was first proposed by Geweke and Porter-Hudak[10] for stationary, fractionally integrated stochastic processes with $d \in (-\frac{1}{2}, \frac{1}{2})$. Later, Hurvich[9] has shown that this approach can also be applied to non-stationary cases of $d \geq \frac{1}{2}$.

Performing a least-squares fit of the model (5) on the average PSD up to the Nyquist frequency of 1.5 mHz has resulted in the following estimate: $d_* = 0.671$. In order to determine the statistical properties of the overall model estimator (including average PSD and fractional power d) for the exactly same case of non-uniform sampling in the 18 observation sequences, a Monte Carlo experiment has been performed. For a time series of the form (4) and $d = d_*$, the distribution of the fractional power estimates (10,000 in total) has a Gaussian shape with a mean of 0.67 and a standard variation of 0.033. Figure 8 shows the histogram of the d-estimates for this experiment.
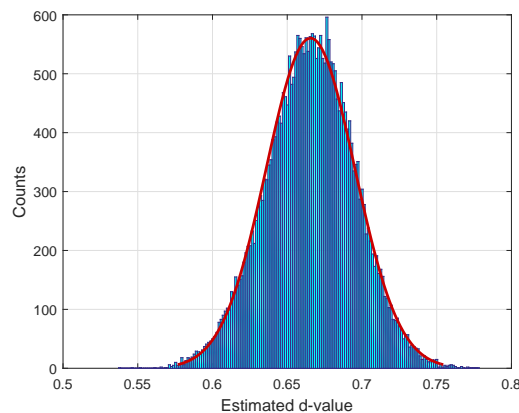

Figure 8. Distribution of d-estimates for 10,000 realisations.

3.2 Residual model error

The fractionally integrated stochastic model with $d_* = 0.671$ captures the fundamental, average spectrum of the dynamic behaviour of the log coherence length. This is clearly demonstrated in Figure 7. To check for additional dynamic behaviour we have taken the residual PSD function and plotted the corresponding auto-correlation function; see Figure 9. This auto-correlation function has a clear peak at lag 0 and besides that a very minor tail. This shape represents a 'nearly-white' residual, indicating that a very significant part of the dynamic behaviour is covered already by the low order, fractionally integrated model.
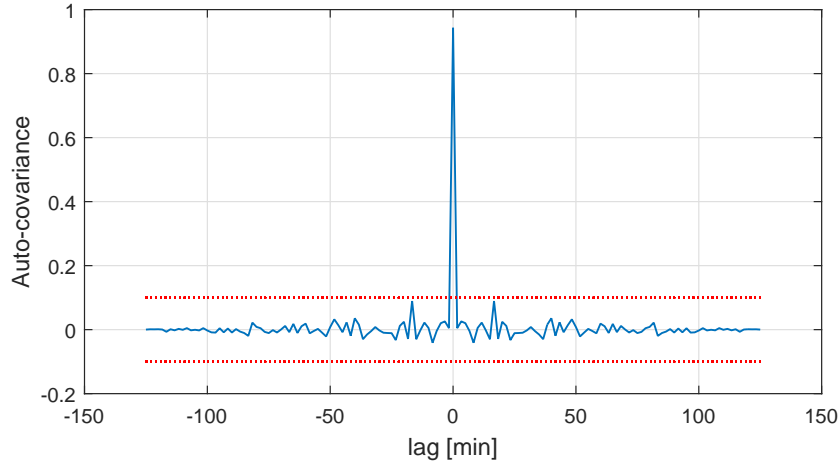

Figure 9. Auto-correlation function of residual model error.

4. FORECASTING OF THE COHERENCE LENGTH

Having a suitable, low-order model for the time variability of the turbulence coherence length gives us the opportunity to address forecasting of r_0. For this particular case of fractionally integrated random noise, in Beran[11] it is shown that the mean square of the prediction error (MSE) over a horizon of k samples has the following characteristic in the limit of $k \to \infty$:

$$MSE(k) \quad \sim var(x_r(k)) \ ... \ \text{for} \ ... \quad d \in (-\frac{1}{2}, \frac{1}{2}), \tag{6}$$

$$MSE(k) \quad \sim \frac{\sigma_\epsilon^2}{2\pi} g(d) k^{2d-1} \ ... \ \text{for} \ ... \quad d \in (\frac{1}{2}, \frac{3}{2}) \tag{7}$$

where in the latter case $g(d) = [\Gamma(3 - 2d) \sin(\pi d)]/[(d - 1)(2d - 1)]$ for $d \neq 1$. This implies that for dynamic model of the turbulence coherence length and $d_* = 0.671$ the prediction error will grow with the horizon k and diverge as $k \to \infty$. To illustrate the prediction of the coherence length, a single realization of an r_0 time series has been generated based on (4), with a mean and variance matching the r_0 distribution from the Stereo-SCIDAR measurements. Figure 10 below shows the r_0 prediction with 68% and 95% confidence intervals.

It follows from Figure 10 that - for this realisation - the prediction uncertainty grows rapidly with the horizon. Only for a short horizon (up to say 5 samples, which equals 500s here), prediction of r_0 has a limited uncertainty margin. An analysis of the average prediction confidence intervals for this case with mean $\overline{r_0} = 0.11m$ confirms this observation; see Figure 11.

5. CONCLUSIONS

A power spectral analysis on the temporal fluctuations of the turbulence coherence length r_0 has been performed. The analysis is fed by 20 measurement nights of the Stereo-SCIDAR profiler installed at the 2.54m Isaac Newton Telescope, La Palma. A single-parameter stochastic model appears to match the spectrum of the temporal fluctuations of the log of the coherence length r_0. This stochastic model suggests a non-stationary character of
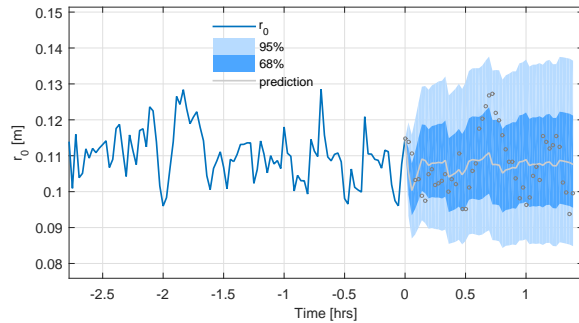
Figure 10. Prediction with confidence intervals for a single realisation of r_0 series. The dotted line shows the continuation of the true signal.
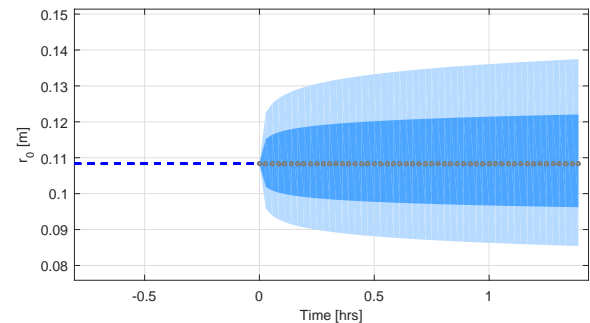

Figure 11. Average prediction confidence intervals (68%, 95%) for coherence length time series with a mean value $\overline{r_0} = 0.11m$.

the r_0 dynamic behaviour. It also suggests that forecasting of r_0 is most beneficial over short horizons; up to 500s ahead in this case. For longer forecast horizons the prediction confidence intervals grow to relatively large values.

ACKNOWLEDGMENTS

REFERENCES

[1] Hardy, J. W., [*Adaptive Optics for Astronomical Telescopes*], Oxford University Press, Oxford (1998).

[2] Ishimaru, A., [*Wave Propagation and Scattering in Random Media*], Academic Press, New York (1978).

[3] various authors in [*Optical Turbulence - Astronomy Meets Meteorology - Conference Proceedings of Optical Turbulence Characterization for Astronomy Applications*], Imperial College Press, London (2009).

[4] Doelman, N., Fraanje, R., and den Breeje, R., "Real-sky adaptive optics experiments on optimal control of tip-tilt modes," in [*Second International Conference on Adaptive Optics for Extremely Large Telescopes*], – (2011).

[5] Shepherd, H. W., Osborn, J., Wilson, R. W., Butterley, T., Avila, R., Dhillon, V. S., and Morris, T. J., "Stereo-scidar: optical turbulence profiling with high sensitivity using a modified SCIDAR instrument," *Monthly Notices of the Royal Astronomical Society* **437 (4)**, 3568–3577 (2014).

[6] Osborn, J., Butterley, T., Fhring, D., and Wilson, R., "Characterising atmospheric optical turbulence using stereo-SCIDAR," *Journal of Physics: Conference Series* **595** (2015).

[7] Racine, R., "Altitude, elevation, and seeing," *Publications of the Astronomical Society of the Pacific* **117**, 401–410 (2005).

[8] Eyer, L. and Bartholdi, P., "Variable stars: Which Nyquist frequency?," *Astronomy and Astrophysics Supplement Series* **135**, 1–3 (1999).

[9] Hurvich, C. and Ray, B., "Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes," *Journal of Time Series Analysis* **16**, 17–41 (1995).

[10] Geweke, J. and Porter-Hudak, S., "The estimation and application of long memory time series models," *Journal of Time Series Analysis* **4**, 221–238 (1983).

[11] Beran, J. and Ocker, D., "Semifar forecasts, with applications to foreign exchange rates," *Journal of Statistical Planning and Inference* **80**, 137–153 (1999).