



Universiteit
Leiden
The Netherlands

Predicting the future: Predictive control for astronomical adaptive optics

Kooten M.A.M. van

Citation

Predicting the future: Predictive control for astronomical adaptive optics. (2020, November 4).
Predicting the future: Predictive control for astronomical adaptive optics. Retrieved from
<https://hdl.handle.net/1887/138080>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/138080>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/138080> holds various files of this Leiden University dissertation.

Author: Kooten, M.A.M. van

Title: Predicting the future: Predictive control for astronomical adaptive optics

Issue Date: 2020-11-04

Chapter 5

Impact of time-variant turbulence behavior on prediction for adaptive optics systems

“However it may prove, one must tread the path that need chooses!”

J.R.R. Tolkien, *The Fellowship of the Ring*

Summary

For high contrast imaging systems, the time delay is one of the major limiting factors for the performance of the extreme adaptive optics (AO) sub-system and, in turn, the final contrast. The time delay is due to the finite time needed to measure the incoming disturbance and then apply the correction. By predicting the behavior of the atmospheric disturbance over the time delay we can in principle achieve a better AO performance. Atmospheric turbulence parameters which determine the wave-front phase fluctuations have time varying behavior. We present a stochastic model for wind speed and model time varying atmospheric turbulence effects using varying wind speed. We test a low-order, data-driven predictor, the linear minimum mean square error predictor, for a near-infrared AO system under varying conditions. Our results show varying wind has a significant impact on the performance of our predictor, preventing it from reaching optimal performance.

5.1 Introduction

Since the first direct image of an exoplanet by Chauvin et al. (2004), high contrast imaging (HCI) has been rapidly advancing. Capable of providing insight into planet formation, characterizing an exoplanet’s atmosphere, and searching for signs of habitability, HCI spatially separates the planet’s photons from the ones emitted by the host star. With multiple dedicated instruments such as SPHERE on the VLT (Beuzit et al., 2008), GPI on Gemini South (Macintosh et al., 2014), and SCExAO on Subaru (Jovanovic et al., 2015) many new observing techniques, better coronagraphs,

and low noise detectors have been developed and tested. These efforts have resulted in processed contrast levels of $10^{-4} - 10^{-5}$ at $0.5'' - 1''$ separation in the near-infrared (Otten et al., 2017). To find cold jupiters, neptunes, and rocky planets, we need improve the contrast levels for HCI systems; aiming for $10^{-7} - 10^{-9}$ at separations of $0.05'' - 1''$ at visible wavelengths.

Currently, one of the most significant limiting factors of a HCI system is the performance of the adaptive optics (AO) sub-systems (Kasper, 2012). Specifically, the time delay (which leads to the servo-lag error) prevents us from observing at small separations as a halo of speckles builds up close to the host star during a science exposure (Milli et al., 2017). In order to improve the performance of the HCI system, see Kasper (2012), the servo-lag error needs to be reduced. One approach to do this is to run the AO system faster, requiring higher speed deformable mirrors, fast wavefront sensors, and bright guide stars, thereby reducing the length of the lag. Alternatively, one can focus on understanding how the phase evolves during the delay and subsequently account for the dynamic behavior during the servo-lag. This is done through predictive control methods, where recent and/or nearby wavefront sensor measurements are used to predict the disturbance over the time lag. (Hardy, 1998)

5.1.1 Predictive control in adaptive optics

Predictive control to mitigate the effect of time delays in AO systems has been worked on for many years (Hardy, 1998). Prediction is not only promising for HCI but also for many different AO applications including improving performance for flux-limited laser guide stars and increasing sky coverage with natural guide stars (Jackson et al., 2015). Linear prediction has been very successful for tip-tilt compensation, especially when induced by structural vibrations as done by Petit et al. (2008), using Linear Quadratic Gaussian control (LQG). Work on the LQG controller for full atmospheric disturbance compensation has also been done by Le Roux et al. (2004). Different predictors, closely related to the Kalman filter, have allowed for filtering in real time and are now being used in laboratory testing as well as in on-sky testing (Paschall and Anderson, 1993; Le Roux et al., 2004; Looze, 2009; Gray et al., 2014; Massioni et al., 2011; Correia et al., 2015). The closely related data-driven H2 optimal controller (Doelman, Fraanje, Houtzager, et al., 2009; Hinnen et al., 2007; Doelman, Fraanje, and Breeje, 2011) was also tested on-sky, for tip/tilt, showing a reduction in the temporal error. More recently, Guyon and Males (2017) show theoretically that Empirical Orthogonal Functions can improve the contrast of a HCI system by a couple orders of magnitude through prediction. Many of these approaches are based on the assumption that atmospheric turbulence induced wavefront phase fluctuations are a stationary stochastic process. However, in nature, turbulence is a time variant process with the mean phase value and variance changing in time. Adaptive predictors have been proposed by van Kooten et al. (2017) to track the time-varying behavior of statistics of atmospheric turbulence. Although some predictors might be robust against non-stationarities, none have been rigorously tested under time-varying conditions. Therefore, in this work, we look at basic data-driven prediction and test

different implementations under time varying conditions.

5.1.2 Atmospheric Non-stationarity

Predictive controllers inherently depend on knowledge of turbulence parameters, as well as the dynamical behavior of the AO system. We know the relevant turbulence parameters such as coherence length, outerscale, and wind speed vary, resulting in variations of the mean and variance of phase aberrations. Atmospheric turbulence induced phase fluctuations are essentially time variant. If we account for this behavior in simulation, we will be able to determine the behavior of a predictor under time varying conditions.

5.1.3 Simulating atmospheric phase

The most recent algorithms for generating atmospheric phase disturbance in simulation, by Assémat et al. (2006), have been proposed to minimize computation time as well as include the effects of boiling (Srinath et al., 2015). They also have flexible architectures that allow for the wind speed, outerscale and Fried parameter to vary at each time step. However, they do not propose how these parameters should vary in time. Without information on how these parameters vary, we are unable to test new control techniques in a simulation environment that reflects all the behavior of the disturbance seen on-sky and instead can only test them under the average behavior.

5.1.4 Outline

In this paper, we focus on understanding the atmospheric wavefront phase fluctuations of an AO system on short time scales and how the disturbance's time-varying behavior affects a data-driven prediction algorithm. In Section 5.2, we look at the dynamical behavior of relevant turbulence parameters—including the wind vector, outerscale, and the Fried parameter—used to describe atmospheric wavefront phase fluctuations. We then present our simulation environment in Section 5.3 and introduce a data-driven prediction algorithm in Section 5.3.2. We present our results for varying wind in Section 5.4 and discuss their impact in Section 5.5. Our future work is proposed in Section 5.6.

5.2 Describing wavefront phase fluctuations

In AO the atmospheric wavefront phase fluctuations are often described as a stationary process with a given variance represented by a spatial covariance function. By assuming Taylor's Frozen Flow hypothesis, we can extend this spatial covariance function to a temporal covariance function through relation with a constant wind speed. Within this framework we can test many different seeing conditions.

Extending the above to time variant wavefront phase fluctuations, we vary the wind speed (v), the Fried parameter (r_0), and outerscale (L_0).

5.2.1 Atmospheric Covariance function

We can describe the spatial covariance of atmospheric phase fluctuation, α , by von Karman model (Conan, 2000) as in Equation 5.1.

$$C_\alpha(r) = \left(\frac{L_0}{r_0}\right)^{5/3} \frac{\Gamma(11/6)}{2^{5/6}\pi^{8/3}} \left(\frac{24\Gamma(6/5)}{5}\right)^{5/6} \left(\frac{2\pi r}{L_0}\right)^{5/6} K_{5/6}(2\pi r/L_0) \quad (5.1)$$

where $\Gamma(x)$ is the gamma function and $K_u(x)$ is the modified Bessel function of order u .

Using the Frozen Flow hypothesis, we can describe a spatial separation, r , as a shift over a short period of time, t , due to v . The temporal covariance follows when this is applied to Equation 5.1.

The wind speed then influences the temporal covariance function. The spatial and temporal covariances are both influenced by L_0 and r_0 . We therefore turn our attention to understanding how these three parameters change in time. First we present a model for the wind vector, then we briefly discuss the behavior of the r_0 and L_0 as they are understood in the above covariance functions.

Wind Vector

Taking wind data from the Thirty Metre Telescope (TMT) site testing campaign at Mauna Kea by Schock et al. (2009), we look at the temporal behavior of the wind speed. The data is from a sonic anemometer at a height of 7 metres above the summit of Mauna Kea, with a variable sampling frequency between 10-60 Hertz. We look at the wind speed and direction fluctuations over a period of three years (2006-2008). From this data set we are able to define a stochastic model that describes the averaged power spectral density (PSD) from night time wind speed data for the ground layer. For each night we look at wind speeds faster than 1 m/s. We observe substantial fluctuations in wind speed (greater than 1 m/s) and wind direction (tens of degrees) at frequencies of 1 Hertz or greater. We observe 3 types of wind behavior from our analysis that represent a large portion of the data (Figure 5.2). In general, over the course of the night either the wind direction or the wind speed can be the dominate source for the fluctuating wind vector. We do not observe nights where both change by large amounts at the same time. Besides these two cases we also have nights where the mean wind speed and wind direction as well as their variances stay the same throughout the night. We consider this to be a realistic Frozen Flow case. These 3 wind behaviors observed is summarized in Table 5.1. We chose to ignore wind direction the remainder of this work due to large fluctuations in direction for slow wind speeds (van Kooten et al., 2018).

To estimate a wind speed PSD we split the night into shorter time segments of 10 minutes. We estimate the PSD using the Lomb-Scargle approach for each segment after de-trending and then take the average to determine the overall PSD for that night. We do this for all nights available, determining an average PSD for the

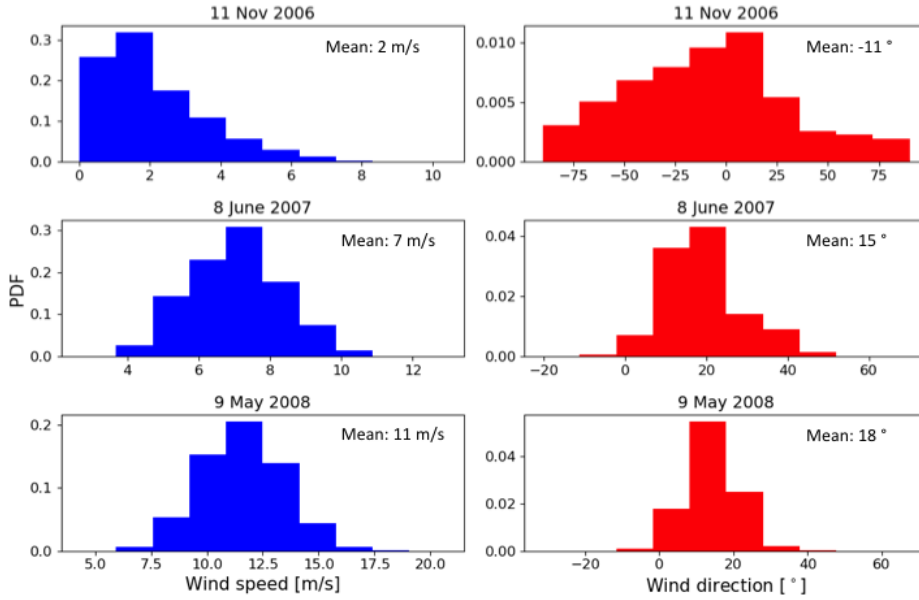


Figure 5.1: Histograms of wind direction and wind speed for three different nights that correspond to the cases in Table 5.1.

Classification	Wind Speed	Wind direction	Example nights
Varying wind direction	Slow (<5 m/s)	Large variance ($\pm 100^\circ$)	2006.11.22
Realistic Frozen Flow	Fast (>8 m/s) with medium variance	Small variance	2008.05.09
Varying wind speed	Large variance (0-20 m/s)	Small variance	2007.06.15

Table 5.1: Three classifications of wind vector behavior seen in our data.

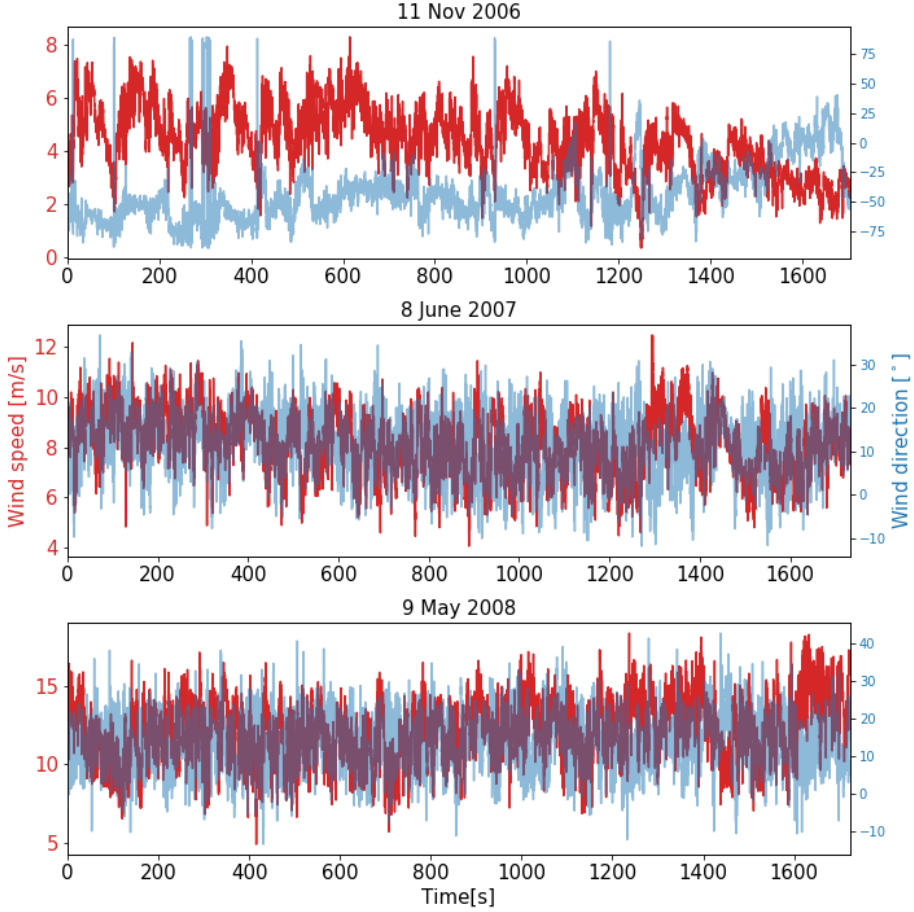


Figure 5.2: Raw time series for three different nights showing the classification in Table 5.1. The data is first filtered using a Butterworth filter to remove a measurement artifact at 20 Hertz due to the sonic anemometer electronics before feeding into our simulations.

wind speed during the period of observations, Figure 5.3. A more in-depth study of the wind statistics and how we estimated the wind's power spectral density is in van Kooten et al. (2018). Here we extend the stochastic model we previously found to capture the roll-off at low frequencies by extending the order of the model proposed to include a tempering component. The resulting stochastic model is a tempered fractionally integrated autoregressive moving average model (ARTFIMA) (Meerschaert and Sabzikar, 2013). Equation 5.2 shows the power spectral density for our model.

$$\Phi(k) = \left| \frac{(1 - 1.0047e^{-ik})}{(1 - 0.999e^{-ik})} \right|^2 |1 - e^{-0.016-ik}|^{-2(0.91)} \quad (5.2)$$

with the angular frequency $k = 2\pi f / f_s$ with f the frequency ranging from 0.002 Hz to 10 Hz and $f_s = 20$ Hz.

We present this model for ground layer wind speed on Mauna Kea for the TMT site location to be used to generate time variant turbulence simulations via varying wind. This model provides a good fit to the individual nightly PSD as well as the averaged PSD, as the overall shape of the PSD remains constant. The gain of the nightly PSD does vary by one order of magnitude throughout the data set.

Fried Parameter

The Fried parameter (r_0) is considered the main turbulence parameter in astronomy that directly influences the achievable imaging resolution. In reality, r_0 varies in time with the turbulence in the atmosphere.

A dynamic r_0 model from Doelman, Fraanje, Houtzager, et al. (2009) was determined from 18 nights of data from La Palma. The model was estimated from data with sampling times ranging from 80 to 120 seconds. This model, therefore, does not include how r_0 behaviors at sub-second timescales and we do not know how it varies on AO timescales, as no other data is available. Therefore, we do not include varying r_0 in our simulations.

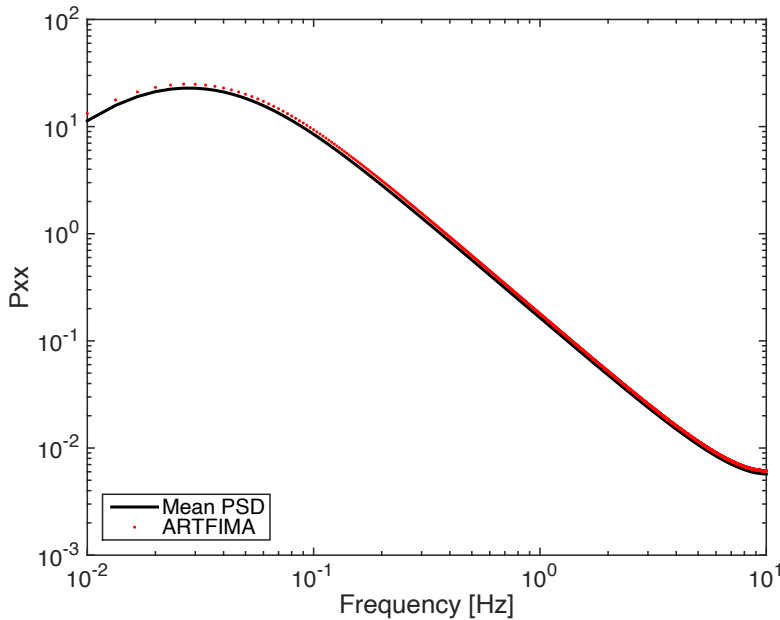


Figure 5.3: Fit of the ARTFIMA model. The black line shows the averaged PSD determined from the data while the red shows the PSD from the ARTFIMA model.

Outerscale

Many efforts have focused on accurately determining L_0 of turbulence with estimations ranging from 5-300 metres (Ziad, 2016). From observational campaigns, average values, probability density functions (PDF), and even a time series over the course of a night have been reported. However, the insights only apply to longer timescales, meaning we do not have sufficient knowledge to model the behavior of L_0 for AO control loop timescales.

5.3 Methods

5.3.1 Turbulence simulation

We use the method developed by Assémat et al. (2006) to generate phase screens with von Karman statistics. It has been included in the HCI Python package, HCIPy being developed at Leiden Observatory by Por et al. (2018). At each time step, the phase screen is updated by the algorithm, allowing for infinite phase screen generation as well as continuous updating of the wind vector (speed and direction), r_0 , and L_0 . This framework allows for non-stationarity behavior to be simulated in an efficient way.

We chose our AO system configuration to be representative of an Extremely Large Telescope (ELT) scale system for HCI based on the planned METIS system for the ELT (Brandl et al., 2016). We model a single layer atmosphere. See Table 5.2 for configuration details. We assume a idealized wavefront sensor where we measure the averaged phase value over the subaperture at a given instantance in time. The total additional noise not due to the servo-lag error has a variance 50 times smaller than the full wavefront phase variance. These choices allow us to focus on the effects of the delay and how a predictor performs.

For the remainder of this work we assume a constant Fried parameter of 13 centimetres at 500 nanometres. For stationary wind cases we use a constant wind speed of 12 m/s unless indicated otherwise. Finally, we assume a constant L_0 of 24 metres unless indicated otherwise.

In our simulations, we predictor using the open-loop phase and estimate the error at each time step for our predictor. We can also choose to predict in closed-loop residual phase. However, in reality, our closed-loop predictor will also depend

Table 5.2: Simulation parameters for our AO system including the sub-aperture (S.A.) size.

λ_{WFS}	D [m]	N_{WFS}	S.A. size [m]	r_0 [m]	Frame delay	Loop speed
2.2 μm	39	74 ²	0.5	0.77	2 samples = 2 ms	1 kHz

on the response of the controller itself and therefore we will need to be account for in the predictor itself.

Wind

Since we have real wind time series available, we use this in our simulations. From the site testing data from TMT, we classify the behavior of the wind vector into three classes as shown in Table 5.1. Six 5 second sections (separated by at least 10 minutes in time) for each case are chosen for three days that are representative cases of our classification. The data are linearly interpolated to the AO loop speed allowing the wind speed to be updated at each time step in the simulation.

5.3.2 Data-driven prediction

In Section 5.2 the time-varying behavior of v , r_0 , and L_0 are described and well as the difficulties in developing a time variant model for atmospheric turbulence. Currently, we only have a model for ground layer v which is site dependent. We, therefore, turn to data-driven methods which use on-sky data to determine the predictor; specifically the linear minimum mean square error (LMMSE) predictor. In this section, we present the framework for the LMMSE predictor and in Section 5.3.5 we place our work in relation to some of the other prediction methods for AO.

We denote the phase at a single point i of a phase screen at time t , $y_i(t)$, and $\vec{u}(t)$ a $P \times 1$ column vector containing a collection of P phase values on a discrete spatial grid at time t (therefore P is the spatial order). We then assume that the future value of a given phase point, $\hat{y}_i$ at the discrete time index $t + d$, is a linear combination of the most recent phase values (our regressors) at time t . Note that we do not have to use all the phase values on our spatial grid but are able to select the spatial order by choosing a region around our point of interest (changing P). We use $\vec{a}_i$ to represent the predictor coefficients. With the above, we define the following where we use T to denote the transpose:

$$\hat{y}_i(t + d) = \vec{a}_i^T \cdot \vec{u}(t) \quad (5.3)$$

$$\vec{u}(t) = (y_0(t) \ y_1(t) \ y_2(t) \ \dots \ y_P(t))^T \quad (5.4)$$

We can expand this to include a set of Q most recent measurements and thereby including more temporal information.

$$\vec{w} = (\vec{u}(t)^T \ \vec{u}(t-1)^T \ \vec{u}(t-2)^T \ \dots \ \vec{u}(t-Q)^T)^T \quad (5.5)$$

We therefore allow for both spatial and temporal regressors through Equation 5.5, where w is a vector of $PQ \times 1$. We have a spatial order of P and a temporal order of Q resulting in the total of PQ regressors.

With $\langle \rangle_t$ being the time average operator, we minimize the cost function:

$$\min_{\vec{a}_i} \langle ||y_i(t+d) - \vec{a}_i^T \vec{w}(t)||^2 \rangle_t \quad (5.6)$$

finding the vector of predictor (filter) coefficients, $\vec{a}_i$, now a $PQ \times 1$ vector. Equation 5.6 represents the time-averaged mean square prediction error.

The strength of this method is that the algorithm immediately extracts any spatio-temporal information from the data - the regressors - provided. It is important to realize for AO that the temporal sampling is much finer than the spatial sampling.

Solving Equation 5.6 for our zero-mean stochastic process, the solution can be written in terms of the inverse of the auto-covariance matrix and cross-covariance vector (Haykin, 2002).

$$\vec{a}_i = \mathbf{C}_{\vec{w}\vec{w}}^+ \vec{c}_{\vec{w}y_i} \quad (5.7)$$

where $+$ denotes a pseudo inverse. $\mathbf{C}_{\vec{w}\vec{w}}$ is the auto-covariance matrix of $\vec{w}$, the vector containing the regressors, and $\vec{c}_{\vec{w}y_i}$ is the vector containing the cross-covariance between the true phase value, y_i and $\vec{w}$.

Looking at Equation 5.7 the optimal AO predictor based on the spatial covariance (Equation 5.1) can be derived for stationary turbulence (Beghi et al., 2007). However, we choose a data-driven approach where we estimate the covariances using measured data collected over a short period. During the collection period (or training period), this batch LMMSE predictor cannot be used (or we can use a previously calculated solution) and is considered off-line.

In the batch-wise approach, once the prediction vector is found it is fixed for the rest of the simulation. By making use of the matrix inversion lemma, we can form a recursive solution in which the pseudo-inverse of the auto-covariance is updated according to Equation 5.8 and the cross-covariance is updated by Equation. 5.9 (Haykin, 2002).

$$\mathbf{C}_{\vec{w}\vec{w}}^+(t) = \mathbf{C}_{\vec{w}\vec{w}}^+(t-1) - \frac{\mathbf{C}_{\vec{w}\vec{w}}^+(t-1)\vec{w}(t-1)\vec{w}^T(t-1)\mathbf{C}_{\vec{w}\vec{w}}^+(t-1)}{1 + \vec{w}^T(t-1)\mathbf{C}_{\vec{w}\vec{w}}^+(t-1)\vec{w}(t-1)} \quad (5.8)$$

$$\vec{c}_{\vec{w}y_i}(t) = \vec{c}_{\vec{w}y_i}(t-1) + y_i(t-1)\vec{w}^T(t-1) \quad (5.9)$$

Substituting Equations 5.8 and 5.9 into Equation 5.7 can then update the prediction vector using:

$$\begin{aligned} \vec{a}_i(t) = & \vec{a}_i(t-1) + \mathbf{C}_{\vec{w}\vec{w}}^+(t-1)y_i(t-1)\vec{w}^T(t-1) - \\ & (\vec{c}_{\vec{w}y_i}(t-1) + y_i(t-1)\vec{w}^T(t-1)) \frac{\mathbf{C}_{\vec{w}\vec{w}}^+(t-1)\vec{w}^T(t-1)\vec{w}(t-1)\mathbf{C}_{\vec{w}\vec{w}}^+(t-1)}{1 + \vec{w}^T(t-1)\mathbf{C}_{\vec{w}\vec{w}}^+(t-1)\vec{w}(t-1)} \end{aligned} \quad (5.10)$$

The recursive solution goes on-line immediately with the initial covariance being set to diagonal matrices with large values (as done with recursive least-squares methods). The solution is quick to converge compared to gradient based methods. The recursive solution depends on all previous measurements finding the optimal

solution for all previous data.

5.3.3 The effect of atmospheric parameters on prediction

As mentioned above, using the covariance in Section 5.2.1 we can derive our optimal predictor in terms of the theoretical covariance. We can then see how the predictor depends on the atmospheric parameters.

For example, the diagonal elements in $\mathbf{C}_{\tilde{w}\tilde{w}}$ (Conan, 2000) are given by:

$$C_\alpha(0) = \frac{L_0}{r_0} \frac{\Gamma(11/6)}{2\pi^{8/3}} \left(\frac{24\Gamma(6/5)}{5} \right)^{5/6} \quad (5.11)$$

while the off-diagonal elements are given by Equation 5.1.

We can then represent $\mathbf{C}_{\tilde{w}\tilde{w}}$ as a matrix $\mathbf{F}_{\tilde{w}\tilde{w}}$ multiplied by the scalar $C_\alpha(0)$. Therefore:

$$\mathbf{C}_{\tilde{w}\tilde{w}}^+ = \frac{1}{C_\alpha(0)} \mathbf{F}_{\tilde{w}\tilde{w}}^+ \quad (5.12)$$

Similarly, we can represent the cross-covariance vector as:

$$\tilde{\mathbf{c}}_{\tilde{w}y_i} = C_\alpha(0) f_{\tilde{w}y_i} \quad (5.13)$$

Inserting Equations 5.13 and 5.12 into Equation 5.7 we obtain the following:

$$\tilde{\mathbf{d}}_i = \mathbf{F}_{\tilde{w}\tilde{w}}^+ f_{\tilde{w}y_i} \quad (5.14)$$

Equation 5.14 does not depend on r_0 (only present in g), but is instead dependent on L_0 and ν (applying Frozen Flow approximation). Therefore, when ν or L_0 vary our optimal predictor changes as well.

5.3.4 Dealing with time-varying behavior

Both the batch and recursive methods are originally designed for stationary problems and so we apply them to our stationary case. To make these methods suitable for time varying cases, the batch can be retrained and reset as frequently as needed. We can introduce a forgetting factor for the recursive LMMSE that is applied during the predictor update, Equation 5.8. This allows for the recursive method to ignore the previous states of the turbulence by gradually reducing the weighting of old observations in the prediction.

5.3.5 Comparison to other predictors

Before looking at the performance and behavior of the LMMSE predictor we first put our work in context of prediction for AO. We aim to give a brief overview of prediction in AO for methods with similar structure to ours but do not claim to be

a complete overview. The batch LMMSE is very similar to the Empirical Orthogonal Function (Guyon and Males, 2017) approach with a slightly different implementation for the matrix inversion. Many other data-driven predictors have been proposed including the traveling wave predictor by Hardy (1998) that makes use of the Frozen Flow hypothesis. By knowing the wind vector, one can spatially shift the current phase measurements; predicting what the wavefront will look like at the time of our correction. We can therefore represent the traveling wave predictor in our own framework where we do not need knowledge of the wind speed but instead it is extracted intrinsically in the LMMSE algorithm. More specifically, if we choose our regressors such that we only allow the most recent measurement, our LMMSE will find a pure spatial solution that is equivalent to the traveling wave predictor. Finally, we can have a pure temporal solution that only uses recent measurements for the phase point of interest. When we limit ourselves to the most recent measurement only for the batch-wise LMMSE predictor, we get the equivalent of the closed-loop integrator with some gain. The integrator and our s1t1 are often not classified as a predictor; however, they are sometimes considered zero order predictors (Hardy, 1998). If we include multiple previous measurements, we can form a higher order temporal predictor that is similar to an AR-structure approach (Le Roux et al., 2004) (that might be used in the optimal control framework).

Within the multi-conjugate AO (MCAO) community, the inclusion of a prediction step combined with phase reconstruction has result in the spatial angular (SA) predictor which makes use of Frozen Flow to perform spatial shift, for a given layer, as part of the reconstruction. Looking at a single layer, on-axis, the SA would perform the same as our spatially only LMMSE algorithm (Correia et al., 2015; Jackson et al., 2015). Other MCAO work has focused on prediction using a Kalman filter as the phase estimator (Piatrou and Roggemann, 2007) and then performing a moving grid prediction (based off prior wind knowledge) during the DM fitting step. This combines both spatial and temporal information to perform prediction which is also possible in our frame work.

5.4 Simulation Experiments

In this section, we present and discuss a number of different simulation results, starting with the case of stationary turbulence in order to gain intuition on how the LMMSE behaves under these conditions. We then present the results for time variant turbulence, applying the LMMSE predictor for the case of varying wind speed from field measurements.

5.4.1 Stationary turbulence

We limit the input data into the LMMSE algorithms by selecting a grid size of 5-by-5 phase points (regressors) of the most recent measurement to predict the central phase point 3 time steps into the future. For ease, we abbreviate our LMMSE with spatial order 5 and temporal order 1 to s5t1 (see Section 5.3.2). We then define our

‘s1t1’ solution as our classical solution. For this case, the algorithm only sees the most recent measurement of one spatial point. This is a pure temporal solution with a history of 1 frame (the most recent measurement). We simply train it with a few frames and keep the solution static. Looking at the pure spatial LMMSE we see the prediction coefficients found by the LMMSE in a Frozen Flow case with a $r_0 = 0.13$ metres (at 500 nanometres) and $v = 10$ m/s, Figure 5.4. Both the batch and recursive solutions show that the phase value at the yellow grid point has been identified as the largest contributor to the behavior of the central phase (outlined in red) at the next time step. Here the phase itself is the largest contributor. However, they both give a slight weighting (~ 0.1) to the neighboring point, indicating that the wind is shifting horizontally. The solution is not spread over many spatial points because the sub-apertures are so large and we are sampling the atmosphere at 1 kiloHertz, with the phase screen traveling less than 0.5 subapertures per sample interval.

On the right in Figure 5.4 we plot the results for the stationary case. We run 10 simulations with the same input parameters (for all future runs as well). We calculate the root-mean-square (RMS) for the 10 residual phase screens at the same time instance. We shaded regions indicate the 1-sigma level. We see that s5t1 performs better than our s1t1 for a single layer Frozen Flow atmosphere. This behavior holds true under a variety of turbulence conditions (different values of v , r_0 , and L_0). This validates our LMMSE algorithms; under stationary conditions, we can use prediction to increase our performance compared to the static pure temporal solution s1t1.

In the first 5 seconds of Figure 5.5 we show how different spatio-temporal predictors with different orders converge differently and are even able to achieve different final performance levels despite having converged fully for different amounts of spatio-temporal information. The amount of data in the solution controls how fast the solution fully converges as well as the residual variance. For higher order spatio-temporal solutions, the convergence rate is slower but the residual variance is smaller. It is important to note that the amount of input data increases much more for increases in spatial information as we radially increase the spatial information (as opposed to temporal information) as to not be biased towards a specific wind direction. The LMMSE is convenient in that it allows us to choose the amount of input data used for prediction to minimize computation time - either through an identification routine (i.e., single value decomposition while finding the batch solution or manually based on prior knowledge). We choose to limit the amount of information by knowledge of the physical constraints of the AO system such as the spatial sampling of the wavefront sensor and the temporal resolution as well as wind conditions. For example, for our system we expect a maximum of 14 m/s winds, therefore, our turbulence in one time step does not move more than one subaperture. Hence, we need one spatial subaperture minimum to capture most of the effects. We chose to increase this to two subapertures to provide edge information but maintaining symmetry as to not have preference to a wind direction means we land up with a 3-by-3 grid instead of the full grid of 74-by-74. Not only is this attractive to reduce computation but for recursive methods, allows for the solution to converge faster.

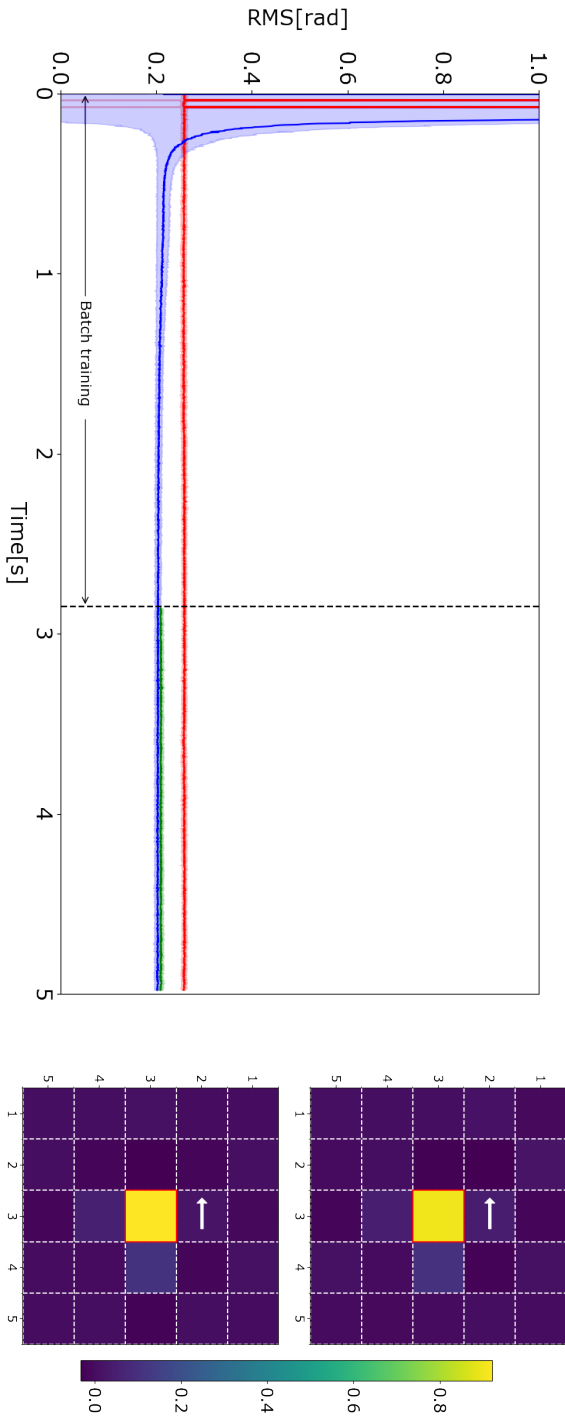


Figure 5.4: Left: The mean residual waveform error for multiple simulations as a function of time for our s1t1 classical solution (red line), batch LMMSE (green line), and recursive LMMSE (blue line). Right: The prediction coefficients for one phase point (outlined in red) for both the batch (top) and recursive (bottom) solutions. The arrow indicates the wind direction.

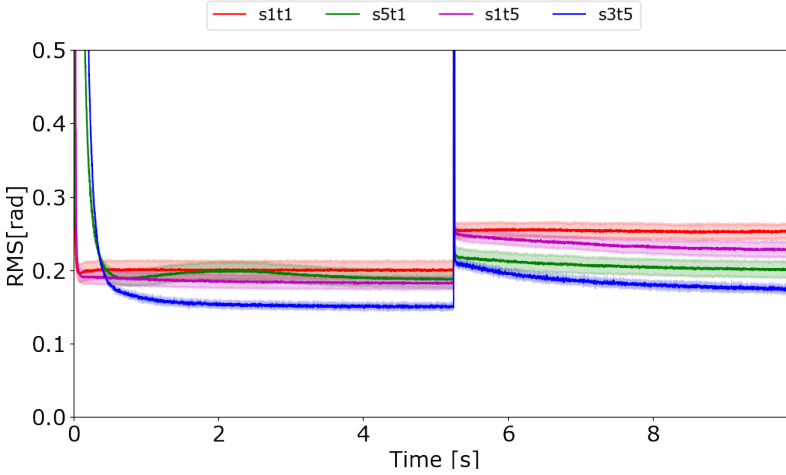


Figure 5.5: Error and convergence of linear predictors with various regressor orders (top) before and after a jump in wind speed (bottom).

5.4.2 Effects due to time varying wind speed

With knowledge of how the LMMSE behaves under stationary conditions, we can study the effects of time varying atmospheric turbulence on our predictors. For this work we set the regressors to be a 3-by-3 spatial grid for 5 previous measurements for each phase point (s3t5). We focus our attention on the effect of a varying wind.

To start, we look at a very simple case for varying wind. Figure 5.5 shows the effect of an instantaneous wind speed jump of 7 m/s. First, we note that as the wind speed increases, the benefit of our predictor increases (i.e., the relative difference between the s3t5 solution and the classical solution). We can see that the recursive algorithm takes a long time (a few seconds) to converge back to its optimal solution after a change in wind speed.

In Figure 5.6 we compare the three different types of LMMSE algorithms for our s3t5 solution under a wind speed jump. The first being the non-adaptive solution: the recursive LMMSE predictor. By knowing the moment when the wind speed changes we are able to recalibrate our batch LMMSE predictor to minimize the loss in performance - this is one type of adaptive solution. Finally, we show our second adaptive algorithm (and third LMMSE algorithm), the forgetting LMMSE predictor. We can see that after the wind jump the forgetting converges faster than the recursive. Eventually the forgetting LMMSE predictor converges the optimal solution calculated off-line in this case. The recursive LMMSE predictor updates the covariance estimate when new data becomes available (Equations 5.8 and 5.9). The estimation makes use of all previous data. Therefore, it will not be able to forget the data before the jump and therefore, will not achieve the same performance as the other algorithms.

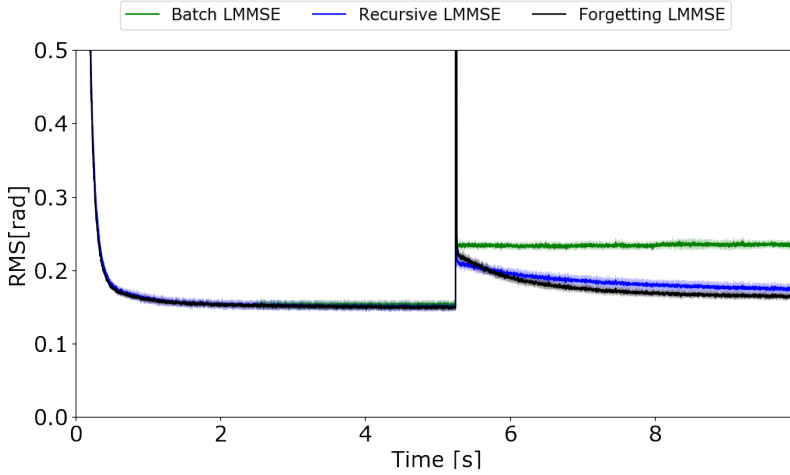


Figure 5.6: Wind jump of 7 m/s on a METIS scaled system showing the convergence of the recursive LMMSE (blue) and the forgetting LMMSE (black) as well as the resetting of the batch LMMSE (green) for s3t5.

Wind speed fluctuations

We now look at the behavior for wind speed fluctuations, shown from Figures 5.7 to 5.12. We do this for two different L_0 values: one smaller than the telescope diameter (24 metres) and one larger than the telescope diameter (80 metres). We plot our adaptive LMMSE solutions: the forgetting LMMSE and the resetting batch LMMSE. As in reality, the batch has no knowledge of the wind speed behavior and is arbitrarily reset. For the classical solution (s1t1), we once again plot the s1t1 solution. Finally, we plot our ideal s3t5 predictor. The ideal predictor is the s3t5 solution under infinite tracking speed and mimicking an instantaneous stationary case. We pre-determine the s3t5 predictor for each potential wind speed assuming stationary conditions and then use a look-up table to find our ideal predictor performance for Figures 5.7 to 5.12. This provides us with a lower bound of a s3t5 predictor's performance. We run the simulation for multiple cases of varying wind on the same night and average the results. This is done for the three different nights shown in Section 5.2.1. We plot the mean wind speed and $1-\sigma$ (standard deviation) levels (shaded region) for the wind speed as well as the different predictors.

5.5 Analysis

We summarize the results of Section 5.4.2 and Figures 5.7 through 5.12 in Table 5.3. We then make the following conclusions:

1. All of our LMMSE approaches lose performance when compared to the ideal predictor under varying wind (ratio for the batch and forgetting LMMSE is greater than 1).

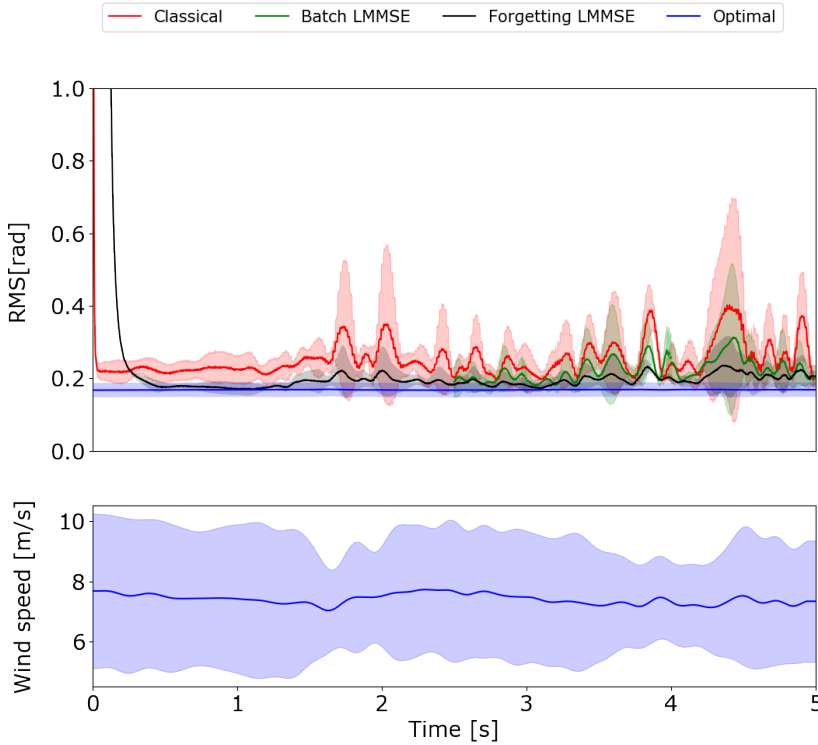
Figure 5.7: 2007.06.15, $L_0 = 24$ m

Table 5.3: Summary of Figures 5.7 to 5.12 showing the ratio of the mean rms value of an algorithm and the mean rms to the optimal LMMSE.

Date	2007.06.15		2008.05.09		2006.11.22	
Outerscale	24 m	80 m	24 m	80 m	24 m	80 m
Classical	1.42	1.40	1.96	1.59	1.19	1.49
Batch LMMSE	1.21	1.19	1.83	1.55	1.59	1.77
Forgetting LMMSE	1.08	1.13	1.47	1.35	1.04	1.36

2. The forgetting implementation does better than the resetting batch implementation (forgetting LMMSE ratio is smaller than the batch LMMSE ratio).
3. The difference between s3t5 and the classical solution becomes smaller under varying wind.
4. The optimal solution has a much smaller variance compared to the other algorithms.

We conclude that a different predictive control approach (that can either track faster or is less sensitive to non-stationarities) is needed to predict atmospheric turbulence in the presence of time-varying turbulence as it is noticeably different than the sta-

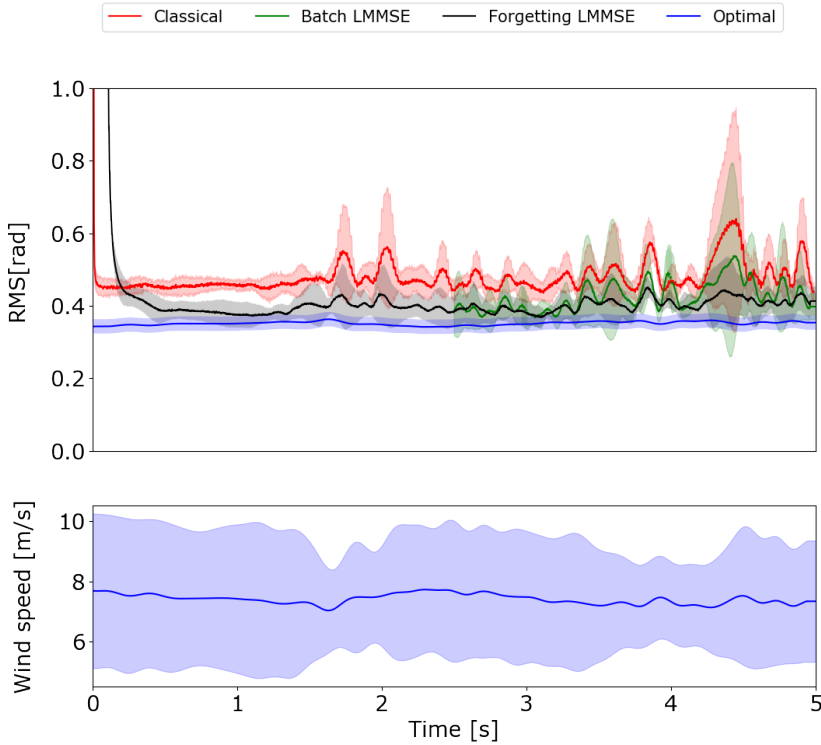


Figure 5.8: The average response for our predictors (upper panel) for three different wind conditions corresponding to Table 1 (lower panel), with $L_0 = 80$ m for wind on the day of 2007.06.15. We plot the s1t1 predictor (red), the s3t5 resetting batch LMMSE (green), the s3t5 forgetting LMMSE (black), and the s3t5 ideal LMMSE (blue). We first allow the batch to train on 2.5 seconds of data (to be consistent with previous simulations) and then reset it every 1 second.

tionary case.

More specifically, from the plots in Figures 5.7 through 5.12 the benefits of our LMMSE predictors are modest. From the optimal solution, we not only see a loss in residual mean wavefront error (see Table 5.3) but also have much larger variances for our LMMSE. Ideally, a predictor would also provide a stable correction.

In all three wind cases, the forgetting LMMSE predictor is able to do fairly well, reaching the ideal solution, for the first couple of seconds. However, the algorithm is unable to handle the non-stationarities as seen by the increase in variance compared to the optimal solution that has infinite tracking capabilities. We attribute this to the continuously changing wind. As we see in Figure 5.6, the forgetting algorithm needs between 0.5 s to 1 s to recover from a large jump, in part due to the number of regressors, and in part due to the magnitude of the change in wind speed. The forgetting LMMSE predictor performs better for slowly evolving changes than fast

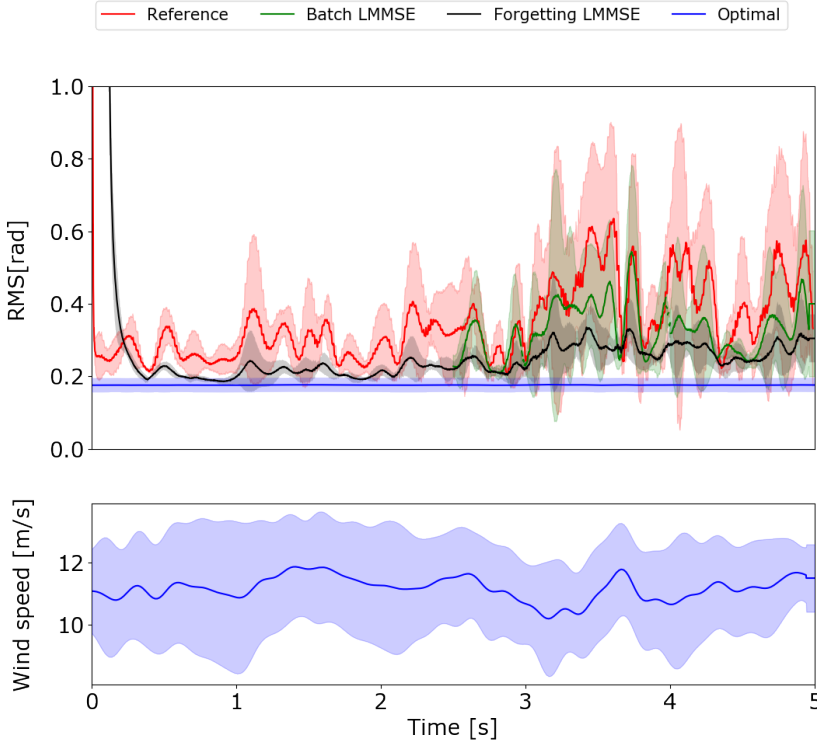


Figure 5.9: The average response for our predictors (upper panel) for three different wind conditions corresponding to Table 1 (lower panel), with $L_0 = 24$ m for wind on the day of 2008.05.09. We plot the s1t1 predictor (red), the s3t5 resetting batch LMMSE (green), the s3t5 forgetting LMMSE (black), and the s3t5 ideal LMMSE (blue). We first allow the batch to train on 2.5 seconds of data (to be consistent with previous simulations) and then reset it every 1 second.

changes. From Table 5.3 we see that the forgetting LMMSE does better than the resetting batch predictor in every case.

5.5.1 Further Discussion

From the stationary tests we can conclude that the LMMSE algorithms with higher order spatio-temporal solutions have a benefit compared to a pure temporal solution making use of only the most recent measurement as we would expect from using Frozen Flow to generate our atmospheric turbulence phase. This agrees with the literature that prediction (whether model-based or data-driven) can provide a reduction in the servo-lag error and the overall residual wavefront error.

The wind fluctuations reported by the TMT site testing group adversely impact the basic predictor—the LMMSE—except for the specific case of slowly moving wind speed. Wind speeds can fluctuate by over 1 m/s on time scales less than 1 second.

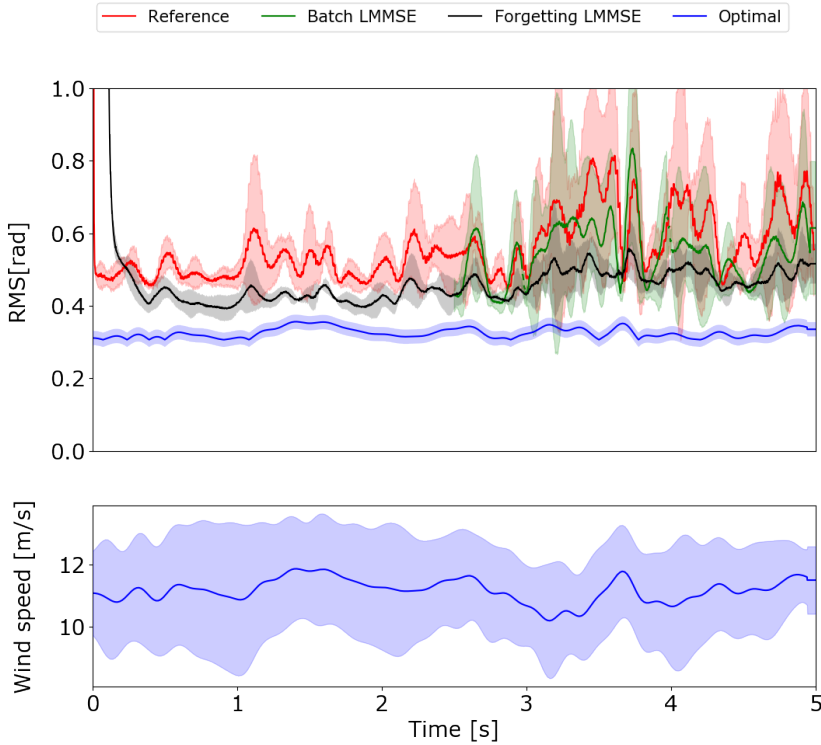


Figure 5.10: Same as Figure 5.7 but for $L_0 = 80$ m with winds according to the date 2008.05.09.

Under time-varying conditions we would expect only a minor performance improvement for a single conjugate AO system with a low order LMMSE predictor compared to no prediction when the system is dominated by the servo-lag error. It is important, however, to note that due to the measurement height of 7 metres these measurements might not reflect what is actually seen by an AO system as the dome and other buildings will influence the low-layer winds. These fluctuations might also be extreme when compared to other astronomical sites. Finally, by limiting our analysis to the ground layer (due to available data), we might be over estimating the strength and timescales of the wind fluctuations and therefore the effect on prediction as in many cases the higher layers are the dominate source for the wind driven halo (Cantalloube et al., 2018).

Assessing our prediction in the presence of our wind fluctuations we quickly see why on-sky test campaigns for predictive AO control are more challenging than common laboratory tests. A single wind jump step already provides a challenge for the LMMSE, moving it away from the ideal predictor solution (Figure 5.5). When we extend our simulation to real wind data, the LMMSE has a harder time tracking the changes (Figures 5.7 through 5.12).

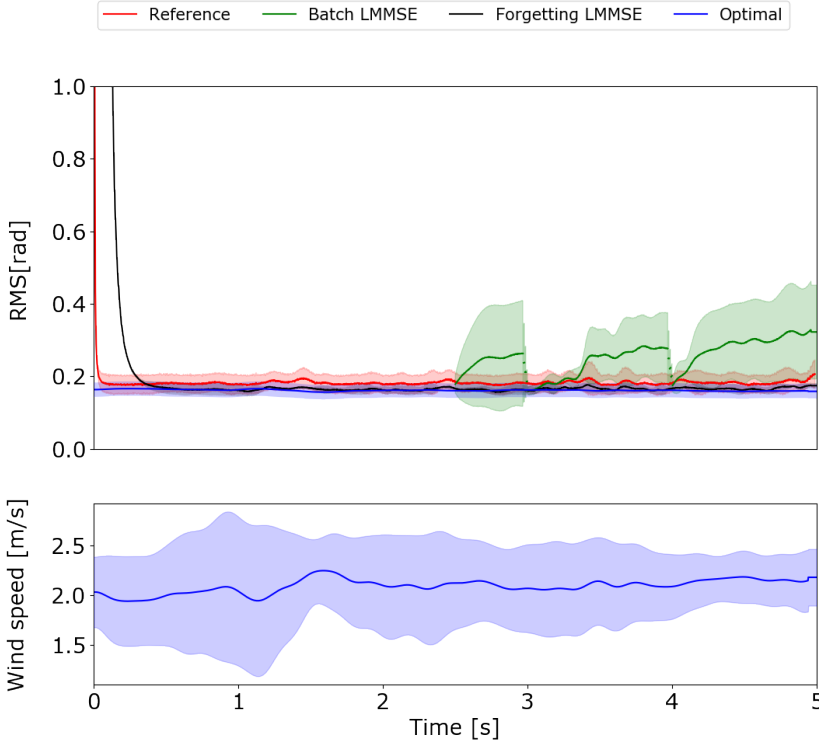


Figure 5.11: Same as Figure 5.7 but for $L_0 = 24$ m for winds according to 2006.11.22.

In the above work we have looked only at the specific case where the temporal delay is a main limiter in performance and hence the use of a LMMSE as a predictor over the time delay only. However, we can also benefit from using these algorithms when we have other dominating error terms. For an increase in a factor of 25 in measurement noise we plot the performance of our high order LMMSE predictors in Figure 5.13. We are able to perform better than the classical solution for varying winds.

5.6 Conclusions

In this paper, we model dynamic wind behavior, based on the TMT site testing data. We look at different implementations of a low order LMMSE predictor to predict atmospheric phase fluctuations over a time delay of two frames. For the stationary case, both the batch and recursive implementations perform better than a pure temporal predictor; i.e., that the current measurement is the best prediction. Under time varying turbulence, however, we loose performance due to wind speed fluctuations. For slower wind speeds and L_0 smaller than the telescope diameter, we are able to approach ideal performance with our LMMSE. The classical solution is also able to

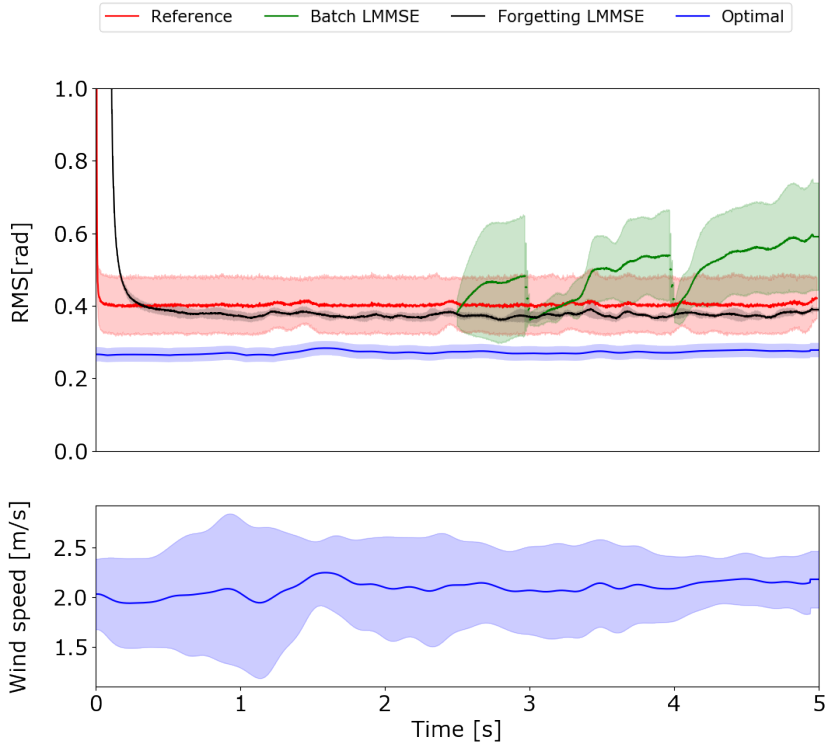


Figure 5.12: Same as Figure 5.7 but for $L_0 = 80$ m with winds according to 2006.11.22.

perform well under these conditions. As the wind speed increases, we see a larger difference between the LMMSE and the ideal solution. Eventually, for high wind speeds, the order of the spatio-temporal predictor no longer affects the performance. The lowest order LMMSE performs just as well as the higher orders. For our given spatial and temporal sampling there are conditions (smaller L_0 and wind speeds approximately m/s) in which we do reach an optimal predictor solution.

This work demonstrates that for an AO system the process of selecting and designing a predictive control framework for time varying wavefront phase fluctuations is much different than for the stationary case. In ongoing research, we are aiming at more suitable predictive control schemes that will be better able to deal with the statistical variability of atmospheric turbulence, especially for cases of high wind speed variance.

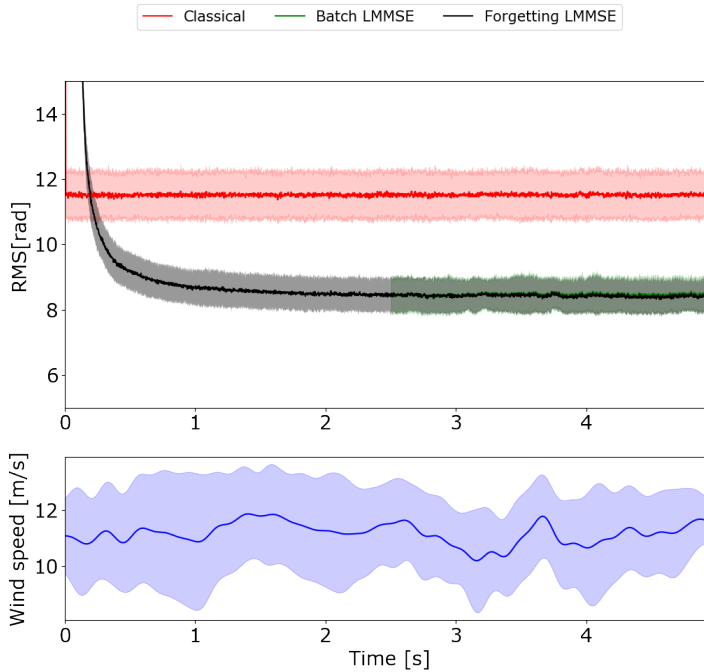


Figure 5.13: Prediction performance for the case of the noise being 0.5 times the turbulence variance.

References

- Assémat, François, Richard W. Wilson, and Eric Gendron (Feb. 2006). “Method for simulating infinitely long and non stationary phase screens with optimized memory storage”. In: *Opt. Express* 14.3, pp. 988–999.
- Beghi, A., A. Cenedese, and A. Masiero (Dec. 2007). “Atmospheric turbulence prediction: A PCA approach”. In: *2007 46th IEEE Conference on Decision and Control*, pp. 566–571.
- Beuzit, J.-L. et al. (July 2008). “SPHERE: a ‘Planet Finder’ instrument for the VLT”. In: *Proc.SPIE* 7014, p. 701418.
- Brandl, B. R. et al. (Aug. 2016). “Status of the mid-infrared E-ELT imager and spectrograph METIS”. In: *Proc.SPIE* 9908, p. 990820.
- Cantalloube, F. et al. (2018). “Origin of the asymmetry of the wind driven halo observed in high-contrast images”. In: *Astronomy & Astrophysics* 620, p. L10.
- Chauvin, G., A.-M. Lagrange, C. Dumas, B. Zuckerman, D. Mouillet, I. Song, J.-L. Beuzit, and P. Lowrance (Oct. 2004). “A giant planet candidate near a young brown dwarf. Direct VLT/NACO observations using IR wavefront sensing”. In: *Astronomy and Astrophysics* 425, pp. L29–L32.

- Conan, R. (2000). *Modelisation des effets de l'échelle externe de coherence spatiale du front d'onde pour l'observation a haute resolution angulaire en astronomie: application a l'optique adaptative, a l'interferometrie et aux tres grands telescopes*.
- Correia, C. M., K. Jackson, J.-P. Véran, D. Andersen, O. Lardière, and C. Bradley (June 2015). "Spatio-angular minimum-variance tomographic controller for multi-object adaptive-optics systems". In: *OSA* 54, p. 5281.
- Doelman, Niek, Rufus Fraanje, and Remco den Breeje (2011). "OPTIMAL CONTROL OF TIP-TILT MODES on-sky adaptive optics demonstration". In: *AO4ELT* 2.
- Doelman, Niek, Rufus Fraanje, Ivo Houtzager, and Michel Verhaegen (2009). "Adaptive and Real-time Optimal Control for Adaptive Optics Systems". In: *European Journal of Control* 15.3, pp. 480–488.
- Gray, M., C. Petit, S. Rodionov, M. Bocquet, L. Bertino, M. Ferrari, and T. Fusco (Aug. 2014). "Local ensemble transform Kalman filter, a fast non-stationary control law for adaptive optics on ELTs: theoretical aspects and first simulation results". In: *Optics Express* 22, p. 20894.
- Guyon, O. and J. Males (July 2017). "Adaptive Optics Predictive Control with Empirical Orthogonal Functions (EOFs)". In: *ArXiv e-prints*.
- Hardy, J. W. (1998). "Adaptive Optics for Astronomical Telescopes". In: *Oxford University Press Jul 1998*.
- Haykin, Simon (2002). *Adaptive filter theory*. 4th. Upper Saddle River, NJ: Prentice Hall.
- Hinnen, Karel, Michel Verhaegen, and Niek Doelman (June 2007). "Exploiting the spatiotemporal correlation in adaptive optics using data-driven H2-optimal control". In: *J. Opt. Soc. Am. A* 24.6, pp. 1714–1725.
- Jackson, Kate, Carlos Correia, Olivier Lardière, Dave Andersen, and Colin Bradley (Jan. 2015). "Linear prediction of atmospheric wave-fronts for tomographic adaptive optics systems: modelling and robustness assessment". In: *Opt. Lett.* 40.2, pp. 143–146.
- Jovanovic, N. et al. (Sept. 2015). "The Subaru Coronagraphic Extreme Adaptive Optics System: Enabling High-Contrast Imaging on Solar-System Scales". In: *Publications of the Astronomical Society of the Pacific* 127, p. 890.
- Kasper, Markus (2012). "Adaptive optics for high contrast imaging". In: *Proc.SPIE* 8447, pp. 8447 - 8447 –10.
- Kooten, Maaike van, Niek Doelman, and Matthew Kenworthy (May 2019). "Impact of time-variant turbulence behavior on prediction for adaptive optics systems". In: *J. Opt. Soc. Am. A* 36.5, pp. 731–740.
- Le Roux, B., R. Ragazzoni, C. Arcidiacono, J.-M. Conan, C. Kulcsar, and H.-F. Raynaud (Oct. 2004). "Kalman-filter-based optimal control law for star-oriented and layer-oriented multiconjugate adaptive optics". In: *Proc.SPIE* 5490, pp. 1336–1346.
- Looze, Douglas P. (Jan. 2009). "Linear-quadratic-Gaussian control for adaptive optics systems using a hybrid model". In: *J. Opt. Soc. Am. A* 26.1, pp. 1–9.
- Macintosh, B. et al. (Sept. 2014). "First light of the Gemini Planet Imager". In: *Proceedings of the National Academy of Science* 111, pp. 12661–12666.

- Massioni, P., C. Kulcsár, H.-F. Raynaud, and J.-M. Conan (Nov. 2011). “Fast computation of an optimal controller for large-scale adaptive optics”. In: *Journal of the Optical Society of America A* 28, p. 2298.
- Meerschaert, Mark M. and Farzad Sabzikar (2013). “Tempered fractional Brownian motion”. In: *Statistics and Probability Letters* 83.10, pp. 2269–2275.
- Milli, J. et al. (Sept. 2017). “Performance of the extreme-AO instrument VLT/SPHERE and dependence on the atmospheric conditions”. In: *AO4ELT5*.
- Otten, G. P. P. L. et al. (Jan. 2017). “On-sky Performance Analysis of the Vector Apodizing Phase Plate Coronagraph on MagAO/Clio2”. In: *The Astrophysical Journal* 834, p. 175.
- Paschall, R. N. and D. J. Anderson (Nov. 1993). “Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements”. In: *OSA* 32, pp. 6347–6358.
- Petit, C., J.-M. Conan, C. Kulcsár, H.-F. Raynaud, and T. Fusco (2008). “First laboratory validation of vibration filtering with LQG control law for Adaptive Optics”. In: *Optics Express* 16, p. 87.
- Piatrou, Piotr and Michael C. Roggemann (Mar. 2007). “Performance study of Kalman filter controller for multiconjugate adaptive optics”. In: *Appl. Opt.* 46.9, pp. 1446–1455.
- Por, Emiel H., Sebastiaan Y. Haffert, Vikram M. Radhakrishnan, David S. Doelman, Maaïke van Kooten, and Steven P. Bos (2018). “High Contrast Imaging for Python (HCIPy): an open-source adaptive optics and coronagraph simulator”. In: *Proc.SPIE* 10703, pp. 10703 - 10703 –14.
- Schock, M. et al. (2009). “Thirty Meter Telescope Site Testing I: Overview”. In: *Publications of the Astronomical Society of the Pacific* 121.878, p. 384.
- Srinath, S., L. A. Poyneer, A. R. Rudy, and S. M. Ammons (Dec. 2015). “Computationally efficient autoregressive method for generating phase screens with frozen flow and turbulence in optical simulations”. In: *Optics Express* 23, p. 33335.
- van Kooten, Maaïke A. M., Niek Doelman, and Matthew Kenworthy (2018). “Implications for contrast as a result of the wind vector and non-stationary turbulence”. In: *Proc.SPIE* 10703, pp. 10703 - 10703 –14.
- van Kooten, Maaïke, Niek Doelman, and Matthew Kenworthy (Sept. 2017). “Performance of AO predictive control in the presence of non-stationary turbulence”. In: *AO4ELT5*.
- Ziad, Aziz (2016). “Review of the outer scale of the atmospheric turbulence”. In: *Proc.SPIE* 9909, pp. 9909 - 9909 –12.

