Reservoir Computing for Prediction of the Spatially-Variant Point Spread Function

Stephen J. Weddell, *Member, IEEE*, and Russell Y. Webb, *Member, IEEE*

Abstract—A new method is presented which provides prediction of the spatially variant point spread function for the restoration of astronomical images, distorted by atmospheric turbulence when viewed using ground-based telescopes. Our approach uses reservoir computing to firstly learn the spatio–temporal evolution of aberrations caused by turbulence, and secondly, predicts the space-varying point spread function (PSF) for application of widely-used deconvolution algorithms, resulting in the restoration of astronomical images. In this article, a reservoir-based, recurrent neural network is used to predict modal aberrations that comprise the spatially variant PSF over a wide field-of-view using a time-series ensemble from multiple reference beacons.

Index Terms—Adaptive optics, reservoir computing, wavefront prediction.

I. INTRODUCTION

LONG exposure imaging through atmospheric turbulence results in the distortion of faint astronomical science objects due to the aberration of effectively planar wave fronts. Wavefront sensors provide indirect measurement of wave fronts from either natural, or artificially created, reference beacons. By measuring the aberrations from a beacon the spatially-invariant (on-axis) PSF is deconvolved with widely used algorithms to restore image quality of a science object, known hereafter as a target object. However, due to the temporal evolution of the atmosphere and corresponding spatial variability of the PSF, the angular separation between a reference beacon and target object is limited by the isoplanatic angle, θ_0. Various methods, such as maximum *a posteriori*-based estimation [1], and modal tomography [2], employ multiple reference beacons for wave front recovery over anisoplanatic regions.

The restoration of astronomical images, distorted by the effects of imaging through the turbulent atmosphere, is based on mathematical principles, applied through digital image and digital signal processing. The real-time restoration of such images can be achieved using adaptive optics (AO), where the optical path of an imaging system is altered by applying the conjugate of the measured wave front to a deformable mirror [3], thus closing the control loop. As AO systems are extended, greater demands are placed on the control system. Open-loop systems, such as deconvolution from wavefront sensing (DFWS) [4], provide the basis for this study by allowing an estimate of a target

object to be made using aberration data from multiple reference beacons. However, to meet real-time requirements for wide field astronomical imaging, an efficient method for the prediction of the spatially-variant PSF is required. In this article we consider a recurrent artificial neural network (RNN) for prediction of modal expansions of a wave front over a wide field-of-view (FOV), employing time-series aberration data from several reference beacons.

This paper is organized as follows. Section II provides a background discussion on several topics in imaging and machine learning that are used throughout this paper. Our approach to restoring images in anisoplanatic regions is outlined in Section III. The use of ANNs in prediction, specifically using reservoir computing architectures, is discussed in Section IV and this is followed by a brief analysis of low-order aberration data in Section V. An outline of our simulation procedures is given in Section VI. The results of our simulations are presented in VII, and real-time considerations of our work are considered in Section VIII. Lastly, our conclusion and outline of future work are given in Section IX.

II. BACKGROUND

In this section we firstly outline atmospheric turbulence and define the wavefront function. Wide-field adaptive optics and wavefront sensing is discussed and this is followed by definitions of the point spread function and image model. An overview of Zernike polynomials, commonly used to represent wavefront aberrations is given, and a brief introduction to artificial neural networks (ANNs) concludes this section.

A. Atmospheric Turbulence

The Earth's atmosphere is comprised of up to three thin layers of turbulence [5]. The movement of large air masses known as weather, is caused by temperature fluctuations and pressure variations. Mixing of air over scale lengths of between one centimetre to several hundreds of meters is the cause of atmospheric turbulence.

The effects of light propagating through atmospheric turbulence can be represented by a wavefront function, $W(x, y)$. A wave front is defined as the optical path-length equivalence (the product of distance, z, and the refractive index, n) from a source object to the pupil plane. In this article only natural guide stars (NGSs) are considered as reference beacons and the region of interest is a wide FOV defined by the entrance pupil, $P(x, y)$, of an optical instrument. The phase fluctuation $\phi(x, y)$ is of primary interest in this article and can be expressed as a function of the wavefront aberration $W(x, y)$ by

$$\phi(x, y) = \frac{2\pi}{\lambda} W(x, y) \qquad (1)$$

where λ is the wavelength and the wave aberration function, $W(x,y)$ is measured along transverse coordinates, (x,y).

B. Wide-Field Adaptive Optics

Wide field adaptive optics is of particular interest in ground-based astronomy. Most target objects are typically too faint for direct wavefront measurements. Single NGSs of sufficient magnitude ($V \leq 14$) provide a wavefront reference, however wavefront errors increase as the angular separation between a target object and NGS exceed the isoplanatic angle, θ_0. Within the visual spectrum, θ_0 is only a few arc seconds [6] and is typically < 40 arc seconds at near infrared wave-lengths. To reduce such limitations, a point-source can be created using a high-power laser and can be positioned within close proximity to a target object. Such reference beacons are referred to as laser guide stars. However, distortions, such as focal anisoplanatism, and the inability to measure tilt aberrations, limit their use as an exclusive alternative to NGSs [7].

Focal isoplanatic errors have been reduced with the use of multiple laser guide stars and application of 3-D tomography. First introduced by Tallon and Foy [8] and supporting a zonal reconstruction method, a modal approach using Zernike polynomials from multiple NGSs has been proposed by Ragazzoni *et al.* [9].

C. Wavefront Sensing & Data Acquisition

Wavefront sensing is an extensive topic discussed in-depth by Roggemann *et al.* [4] and Roddier [3]. In this paper, we limit our discussion to wavefront sensing using the curvature sensing method [3] for two reasons. Firstly, since an estimate of the wavefront phase is measured from the focal, rather than the pupil plane, the curvature method can be used to simultaneously estimate the wavefronts of multiple reference beacons using a wide field charge coupled device (CCD). Secondly, estimating phase from curvature is a nonlinear operation [10] that can be facilitated by the application of a trained ANN [11]. Efficient solutions that employ dimensionally reduction have demonstrated the feasibility of real-time wavefront estimation from focal plane images using a radial basis function neural network [12].

Wavefront data comprising time-series ensembles of Zernike coefficient terms were acquired from field observations using the method described here. These acquired data ensembles form the basis of our simulation models discussed in Section VI-B. We present a time-series analysis of these data ensembles in Section V.

D. The Point Spread Function and the Image Model

In optical signal processing the response of a system to a point source of light is a 2-D impulse response function commonly referred to as a point spread function (PSF). Since an object, $f(\cdot)$, can be represented as a set of weighted point sources, the response is that of a linear system, observing that object, $f(\cdot)$, is the superposition of weighted and shifted versions of the PSF.

Given the coordinate systems (x,y) representing image space and (x',y') representing object space, the PSF can be written, $h(x,y;x',y')$. If the object being observed is $f(x',y')$, the image recorded can be written

$$d(x,y) = \int \int f(x',y') h(x,y;x',y') dx' dy' + \eta(x,y) \quad (2)$$

where $\eta(x,y)$ represents the inevitable presence of noise in the process. This then represents a continuous model for the imaging process.

Without loss of generality and recognising that the image sensors, such as CCDs, are inevitably discrete in nature, a discrete equivalent to (2) can be written

$$d(p,q) = \sum_{k=1}^{N} \sum_{l=1}^{M} f(k,l) h(p,q;k,l) + \eta(p,q) \quad (3)$$

where (p,q) and (k,l) are the indices of discrete coordinates in image and object spaces, respectively. The noise sources $\eta(p,q)$ comprise CCD sensor read noise and photon noise. These noise sources are modelled using Gaussian and Poisson distributions [13], respectively.

In general, $h(p,q;k,l)$ is dependent both on (p,q) and (k,l), i.e., variations of the PSF depend on the location of the source point. In a significant number of imaging situations, however, a spatially invariant model can be adopted, such that (3) becomes a convolution, thus

$$d(p,q) = \sum_{k=1}^{N} \sum_{l=1}^{M} f(k,l) h(p-k;q-l) + \eta(p,q). \quad (4)$$

The optical transfer function (OTF), represented as $H(u,v)$, is the normalized autocorrelation of the pupil function, $P(x,y)$, or simply, the Fourier transform of the PSF. The OTF is analogous to the frequency response of a time-domain filter.

Atmospheric turbulence alters the phase of effectively planar wave fronts, as defined in Section II-A. Such phase alterations over the pupil plane results in distortions in the image plane; such distortions are modelled by the deformation of the PSF.

The relationship between the wave front and PSF is given as [14]

$$h(x,y) = \frac{A_p}{\lambda^2 d^2} \left\| FT \left\{ P(x,y) \exp \left[-j \frac{2\pi}{\lambda} W(x,y) \right] \right\} \right\|^2 \quad (5)$$

where FT is the Fourier transform operator, d is the distance from the exit pupil to the image, A_p is the area of the exit pupil, $P(x,y)$ is the exit pupil function, λ is the wavelength, and $W(x,y)$ is the wavefront aberration at the exit pupil.

Equation (5) relates the PSF to the wave front in the pupil plane, however it does not provide insight into the effects of the PSF on specific wavefront aberrations. For example, a low-order aberration, such as tilt, will result in a displacement of the PSF, whereas higher order aberrations, such as defocus and astigmatism, will result in the deformation of the PSF.

In summary, the temporal effect of air turbulence results in continuous wavefront aberration in the pupil plane and corresponding distortion or shifting of the PSF. The resulting image
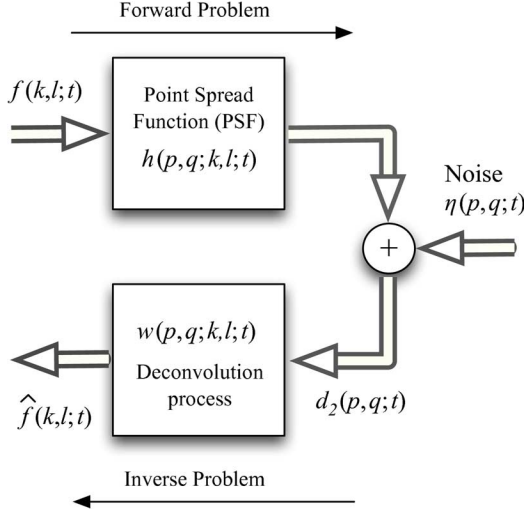
Fig. 1. Generalized imaging model.

is thus formed as a distorted or shifted PSF of the original object, with additive noise due to data acquisition. These aspects are related to the temporal imaging model and are discussed in the following section.

E. Temporal Extensions to the Image Model

In developing an imaging model we need to recognise that the quantities involved are time variant; the object and point spread function and therefore the image also, are all functions of time. For completeness therefore, we should write

$$d_1(p,q,t) = \sum_{k=1}^{N}\sum_{l=1}^{M} f(k,l;t)h(p-k;q-l;t) + \eta(p,q;t) \quad (6)$$

for the time-variant, spatially-invariant image model, and

$$d_2(p,q,t) = \sum_{k=1}^{N}\sum_{l=1}^{M} f(k,l;t)h(p,q;k,l;t) + \eta(p,q;t) \quad (7)$$

for the time-variant, spatially-variant image model.

The image model used to represent the work outlined in this paper is shown in Fig. 1. The input–output relationship of an object, $f(k,l;t)$, can be expressed as distortion and deconvolution operations by respective functions $h(p,q;k,l,t)$ and $w(p,q;k,l;t)$, with additive noise source, $\eta(p,q;t)$. The distortion of object, $f(k,l;t)$, with an aberration represented by the PSF, $h(p,q;k,l;t)$, results in measured data in the spatio–temporal domain, $d_2(p,q;t)$, with additive noise component, $\eta(p,q;t)$. Models of this kind are known as the *forward problem* [15].

The recovery of the original image, $f(k,l;t)$, in the presence of noise, $\eta(p,q;t)$, results in an estimation of the original image, $\hat{f}(k,l;t)$. Restoration problems such as this are referred to as the *inverse problem* and are typically ill-posed.

On-axis measurements using reference beacons allow isoplanatic measurements to be performed, however deviation from on-axis measurements of only a few arc seconds in the visible spectrum [6] result in a decrease in the correlation of Zernike coefficients, as a function of angular separation between a reference beacon and target object.

Prediction of the PSF in anisoplanatic AO images for restoration has been conducted, firstly using wave-optics simulation to compute the PSF for various field angles, and secondly, by fitting each simulated PSF to a parameterized model [16]. The coefficients of the model are then interpolated for prediction of the PSF over a wide range of field angles. Fusco *et al.* [17] showed that the optical transfer function (OTF) could be separated into two components, an OTF that provided the *on-axis* isoplanatic OTF, and a second OTF that defines the *off-axis* anisoplanatic component. By using the correlation of the Zernike coefficients and estimation of the C_n^2 atmospheric profile, the OTF could be theoretically computed over the entire FOV, assuming an infinite exposure time.

Prior knowledge, such as C_n^2 profiles for individual turbulence layers, has been used to minimise the PSF variations over wide angular separations. This technique, known as modal tomography, [9] has been used to estimate modal aberrations, in terms of Zernike polynomials, of anisoplanatic regions over a wide FOV. Modal representation of wavefront aberrations using Zernike polynomials is highly relevant to our work and is discussed in Section II-F.

Since our ultimate objective is in the restoration of images over anisoplanatic regions, our interest is in the spatially variant PSF given by (7). As shown in the preceding discussion and model given in Fig. 1, our analysis is also based on the temporal evolution of the PSF. Justification for this consideration is based on the evolution of the PSF over a wide field-of-view, and thus, the PSF should be considered a function of both time and space. This is further discussed in Section VI-B.

F. Zernike Polynomials

Zernike polynomials are 2-D orthonormal basis functions commonly used to define optical aberrations over the unit circle. They represent the statistical eigenfunctions of optical distortions that quantitatively classify each aberration using a set of polynomials. This set of polynomials is defined as

$$\left.\begin{aligned} Z_{i=even}(r,\theta) &= \sqrt{n+1}R_n^m(r)\sqrt{2}\cos(m\theta) \\ Z_{i=odd}(r,\theta) &= \sqrt{n+1}R_n^m(r)\sqrt{2}\sin(m\theta) \end{aligned}\right\} m \neq 0 \quad (8)$$

$$Z_i(r) = R_n^0(\rho)\sqrt{2}, m = 0 \quad (9)$$

where r is the aperture radius, i is a single index numbering scheme adopted by Noll [18], and the terms m and n are the azimutal and radial order, respectively. The radial polynomial function, $R_n^m(r)$, is defined as

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^n(n-s)!}{s!l\left[\frac{(n+m)}{2}-sl\right]!l\left[\frac{(n-m)}{2}-sl\right]!}\rho^{(n-2s)}. \quad (10)$$

Each Zernike mode can be represented by a 2-D image, commonly referred to as a *phase map*. For example, an aberrated PSF is represented by K phase maps on the pupil plane.

The linear combination of K aberrations over a unit circle of radius R results in an approximation of the phase perturbation

$$\phi(R\rho,\theta) \approx \sum_{i=2}^{K} a_i Z_i(\rho,\theta) \quad (11)$$

where ρ is the normalized aperture given an aperture of radius, R, and when $K = \infty$, (11) is an exact representation of the phase. Additionally, the piston term, Z_1, has been removed in (11), as is common for single aperture instruments.

The Zernike coefficients are defined as

$$a_i = \int_0^1 \int_0^{2\pi} \mathcal{W}(\rho)\phi(R\rho,\theta)Z_i(\rho,\theta)\,d\theta\,d\rho \qquad (12)$$

where $\mathcal{W}(\rho)$ is the pupil weighting function.

The generalized pupil function can be defined as [4],

$$W(x,y) = \mathcal{W}(x,y)\exp[j\phi(x,y)] \qquad (13)$$

where the pupil weighting function, $\mathcal{W}(x,y)$, is more conveniently expressed in rectangular coordinates.

Equation (13) is a modal description of the wave front over the exit pupil of an imaging system. However, Ragazzoni showed that a smaller, circular portion of a wave front on the pupil can be described by another ensemble of Zernike coefficients [2]. We apply this important result by using a modal representation to describe wavefront aberrations in both isoplanatic and anisoplanatic regions for predicting the spatially variant PSF over a wide FOV.

G. Artificial Neural Networks

Artificial neural networks (ANNs) are inspired by biological neurons and their interconnections that form the structure of the brain. ANNs can be defined as a network of nodal structures each comprising one or more input and output nodes, a weighted matrix of interconnections providing communication between each input and output, and an activation function that ensures each output is bounded. When an analytical model is either not available or is too complex, *black-box* modelling can be applied using ANNs to find the solution space. In this regard, ANNs do not necessarily model internal mechanisms but are trained to learn the overall behavior of a target system [19].

A class of ANNs known as recurrent networks supports an architecture based on signal feedback to enhance a systems ability to predict or forecast some future value [20]. Such networks can be employed to capture the dynamics of the input signals. The integration of ANNs and associated topologies for AO applications has augmented or replaced many existing models by providing a simplified methodology [11].

The application of ANNs in AO systems can be grouped into two categories; firstly, the classification of wavefront aberrations in the pupil plane, and secondly, the prediction of wavefront aberrations [11]. Recently, an ANN system architecture has been proposed for the prediction of wavefront aberrations using multiple NGSs [21].

III. Approach

Our approach is to use a recurrent artificial neural network (ANN) to predict the space variant PSF that represents a target object in an anisoplanatic region over a wide FOV. Firstly, a reservoir-based ANN is used to learn the temporal evolution of aberrations from N source objects within an anisoplanatic field, with respect to a nominated target object. Secondly, once trained, the ANN is used to estimate aberrations from a target object, given a generalized distortion field. Lastly, the predicted space variant PSF is compared with the actual space invariant PSF and the resulting mean-squared error (MSE) is used to validate this method.

Rather than employ or estimate a model of the distortion through maximum likelihood, our goal is to estimate a generalized solution. We address the inverse problem of image restoration by reconstructing the spatially variant PSF with simulated aberration data and the use of a trained, recurrent ANN.

Our method used to reconstruct the spatially-variant PSF can be summarized in six stages.

- Stage 1: Determine N isoplanatic regions over a wide FOV; each region is defined by a single NGS.
- Stage 2: Estimate the set of wavefront aberrations from N NGSs over a time-series and the angular separations that exist between each NGS and a nominated target, T. An estimate of the wavefront aberrations, $S_{1...N}^{Z=1...20}$, are incorporated with corresponding angular separations, $\theta_{1...N}^{Ext}$, to form a data ensemble for use as a training set.
- Stage 3: An RNN is trained using data from simulations performed during Stage 2.
- Stage 4: The performance of the RNN is evaluated using test ensembles, exclusive of the training set.
- Stage 5: The RNN is optimized to achieve a minimum MSE over the set of isoplanatic regions that comprise the FOV. Repeat stages 1 to 5 until acceptable performance is obtained.
- Stage 6: The predicted PSF is used with a deconvolution algorithm to reconstruct a portion of the image, supported within a corresponding anisoplanatic region.

Our hypothesis is that by using the temporal correlation properties of the optical field to learn the temporal properties of the index of refraction fluctuations, a recurrent ANN can be employed for prediction of the SVPSF over a wide field, given N NGSs.

Until recently, the prediction of high-dimensional time-series data to within acceptable performance criteria, required specialized, computationally intensive training algorithms. Fortunately, a relatively recent RNN architecture, generally referred in the literature as *reservoir computing*, minimizes the amount of training required [22], and this feature has been fully utilized in our application.

IV. Reservoir Computing

Artificial neural networks (ANNs) that employ supervised learning require training data to provide a generalized solution in K dimensional space. Spatio–temporal data comprising wavefront aberrations from each source, $S_{1...N}$, and separations of a target to multiple source objects are used to train an artificial neural network for predicting the effects of turbulence in anisoplanatic regions.

Specialized, recurrent ANN architectures that simplify the training of ANNs and provide good performance, in terms of generalization and approximation, have been reported in the literature [19]. The generic term, *reservoir computing*, has been given to such recurrent networks and includes echo state networks (ESNs) and liquid state machines. A method will be de-
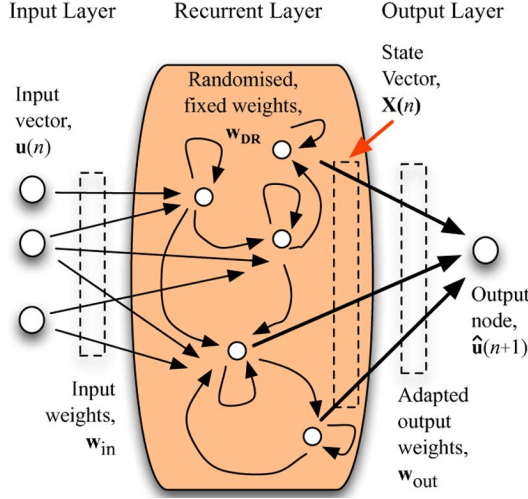
Fig. 2.   Echo State Network (ESN) Architecture.

scribed in this section that utilizes an echo state network (ESN) to predict phase distortions represented by multiple PSFs over a wide FOV. A method using a time-delayed neural network for prediction of aberrations caused by turbulence has been proposed [23]. However, in our view, the complexities in training and implementing such networks can be avoided with the use of ESNs. An extension to the method outlined in this paper for general image processing applications is considered.

A. Architecture

Recurrent networks support a memory structure and are therefore suitable for time-series prediction. ESNs are recurrent networks that simplify training and provide a *reservoir of rich dynamics* [19]. A fixed sparse matrix, $\mathbf{w}_{DR}$, is used to implement the recurrent network and a linear readout, $\mathbf{w}_{out}$, is trained to produce an output. Equation (14) defines the state vector, $\mathbf{X}(n)$, and output vector, $\mathbf{Y}(n)$, of the ESN as shown in Fig. 2:

$$\mathbf{X}(n) = \varphi^{trn}\Big(\mathbf{w}_{in}\mathbf{u}(n)^T + \mathbf{w}_{DR}\mathbf{X}(n-1)^T$$
$$+ \mathbf{w}_{back}\mathbf{Y}(n-1)^T\Big),$$
$$\mathbf{Y}(n) = \varphi^{out}(\mathbf{w}_{out}\mathbf{X}(n)^T) \tag{14}$$

where $\mathbf{w}_{in}$ is the input weight matrix, $\mathbf{w}_{DR}$ is the dynamic reservoir matrix, $\mathbf{w}_{back}$ is the feedback matrix, $\varphi^{(\cdot)}$ are sigmoidal activation functions, and $\mathbf{w}_{out}$ is the output weight matrix.

The predicted output of the input series, $\mathbf{u}(n)$, one time-step into the future, is given by

$$\hat{\mathbf{u}}(n+1) = \mathbf{Y}(n). \tag{15}$$

The ESN is used as a modal predictor to determine Zernike coefficients, given spatio–temporal input data. Zernike terms $Z_{2...20}$ are extracted from each image frame, shown as source reference $S_{1...n}$ as shown in Fig. 5. A resulting time-series ensemble is created over successive image captures and is defined by (18) in Section V-B. The analysis of these resulting time-series is discussed in Section V.

B. Training

In addition to Zernike coefficients, the angular separations between each source, $S_{1...N}$, and the target object, T, are used to train the ESN over a wide spatial field. The projections of each separation on the image plane equate the Euclidian distances between $S_{1...N}$ and T. We denote the separation data as, $\theta_{1...N}^{Ext}$, where N is the number of source objects, and is represented in vector notation as $\mathbf{E}$.

Batch training is used as distinct from *online* training, due to the lower MSE achieved by the former. The recurrent layer is also known as the dynamic reservoir. Randomized, fixed weights associated with the dynamic reservoir are shown in Fig. 2 and are represented in matrix form as, $\mathbf{w}_{DR}$. At each training step the outputs of the dynamic reservoir that form an N-dimensional subspace are used in conjunction with the desired response, $\hat{\mathbf{u}}(n+1)$. The resulting Euclidian distance is used to update the linear coefficients, $\mathbf{w}_{out}$.

Thus, the output weights, shown as $\mathbf{w}_{out}$ in Fig. 2 are trained in this process, however unlike other RNNs that employ the backpropagation through time (BPTT) algorithm [20], the recurrent layer in echo state networks is not trained and remains fixed. The Moore–Penrose *pseudo-inverse* is used to minimise output weights and is given as

$$\mathbf{w}_{out} = (\mathbf{X}.\mathbf{w}_{train})^+ \tag{16}$$

where $\mathbf{X}$ is the state vector, and $\mathbf{w}_{train}$ is the training vector that comprises the Zernike terms $Z_{2...20}$ and spatial data, $\mathbf{E}$.

V. TIME-SERIES ANALYSIS

Time-series analysis was conducted on wavefront aberration data acquired from both field observations and simulation. A comparative study was conducted on these data-sets, firstly to validate our simulations in wavefront propagation and imaging through turbulence, and secondly, to provide an insight into the chaotic behavior of low-order tilt distortions to aid in the prediction of wavefront aberrations. A summary of these results is presented in this section. For a more rigorous temporal analysis of wavefront tilt data see McGaughey *et al.* [24].

A. Field Data Acquisition

Two CCD cameras were used to acquire intra-focal and extra-focal images of two NGSs, $S_{1,2}$, over a 56 arc second FOV. Both cameras were synchronized to capture an image every 16.7 ms. The resulting frame rate of 60 frames/s (FPS) was based on the temporal decorrelation time of the atmosphere, given by the Greenwood frequency [25]

$$f_G = 0.43[V_{wind} \cdot (r_0)^{-1}] \tag{17}$$

where the wind velocity, V_{wind}, was estimated at $2.4 \ \mathrm{ms}^{-1}$ and Fried's coherence length, r_0, was measured at 0.05 m using generalized scintillation detection and ranging (SCIDAR) [26] instrumentation at the Mt. John Observatory, New Zealand. Oversampling was also deemed necessary due to noise considerations and variability of wind speed.

A region of interest was defined for each NGS and the Poisson equation was solved to retrieve wavefront maps using the cur-
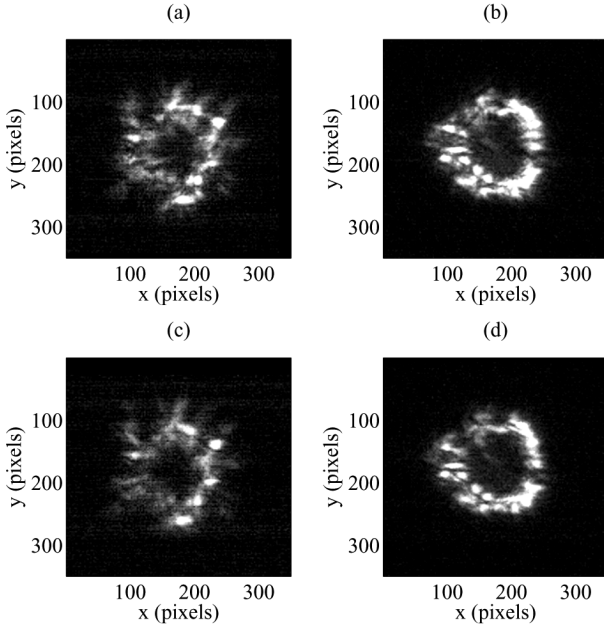
Fig. 3. Short-exposure intra- and extra-focal images of μ^1 Crucis (a) and (b), and μ^2 Crucis (c) and (d), respectively.

Fig. 4. Time-series Zernike aberrations: (a) Defocus (Z_4) and (b) astigmatism (Z_5).

vature method proposed by Roddier [3]. The orthonormal set of Zernike polynomials is used to represent the wavefront aberrations of each NGS.

Data ensembles were created using long exposures, each comprised from a set of 1000 short exposure images. Time-series data defining Zernike coefficients, $a_2 \ldots a_{20}$, representing the wavefront aberrations from each NGS were used as the basis for analysis. To verify our simulation model, short-exposure intra- and extra-focal images were acquired using the wavefront retrieval method outlined previously. As a reference beacon, the binary star μ Crucis was selected. This proved an ideal candidate for our study with a separation of μ^1 and μ^2 Crucis of 35 arc seconds and magnitude, $V \approx 4$ and $V \approx 5.2$, respectively. Typical examples of respective images used for wavefront retrieval are shown in Fig. 3. The analysis of tilt aberrations $Z_{2,3}$ was performed, however higher-order aberrations $Z_{4...7}$ that distort, rather than displace the PSF as outlined in Section II-D, are presented in Section V-B.

B. Analysis

Given N samples, a time series can be represented as

$$\boldsymbol{u}(n) = [u(n), u(n-1), u(n-2), \ldots u(n-N+1)]. \quad (18)$$

Nineteen Zernike coefficients were recorded and time series analysis was conducted on selected aberrations to determine the extent of low-dimensional chaotic attractors and prediction error. Examples of time-series data, such as defocus, Z_4, and astigmatism, Z_5, are shown in Fig. 4.

The following parameters were measured for time-series analysis.

1) The correlation time, τ_c
2) The minimum embedding dimension, D_E.
3) The largest Lyapunov exponent, λ_L.
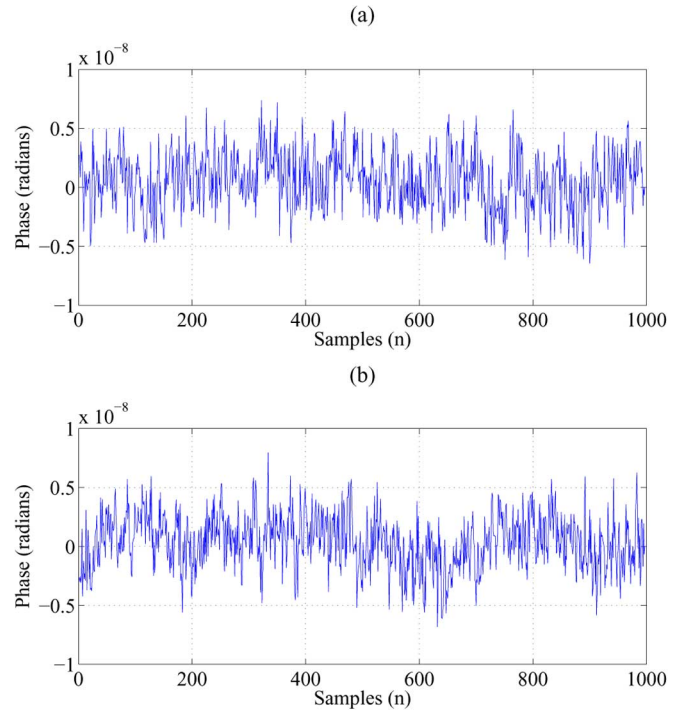4) Correlation dimension, D_r.

TABLE I
TIME-SERIES ANALYSIS—PARAMETER SUMMARY

Time-Series Parameter	Zernike Polynomial		
	Z_4^{Obs}	Z_5^{Obs}	Z_7^{Obs}
Correlation Time, τ_c	2.1	1.9	2.4
Embedding Dimension, D_E	5	4	5
Lyapunov Exponent, λ_L	1.8	2.2	1.85
Correlation Dimension, D_r	2.92	3.9	3.1

The minimum embedding dimension was determined by Cao's method [27] and was employed using the nonlinear time series analysis package, TSTOOL [28].

To estimate the Lyapunov exponents, an algorithm similar to Wolf [29] was used. The average exponential growth of the distance of neighboring orbits was calculated from the prediction error. An estimate of the largest Lyapunov exponents was based on increases in the prediction error, as a function of prediction time. The results of a temporal analysis conducted on low-order wavefront aberrations from observational (field) time-series data is shown in Table I. Based on these results our simulation data was validated and provided a solid basis for further exploration in terms of the optimization of neural network architectures for the prediction of wavefront aberrations over a wide FOV. For example, *a posteriori* knowledge of the chaotic temporal behavior of the system can be used as a basis to establish spectral density and sparseness of the reservoir matrix.

Since the largest Lyapunov exponent is positive, the time series is considered chaotic [30]. The correlation dimension tests the presence of deterministic chaos and has previously been performed on wavefront tilt data [24]. Thus, this analysis implies

that each modal aberration has a deterministic component that can be predicted. Several studies on the prediction of wavefront tilt have been undertaken [12], [31], [32].

C. Modelling Chaotic Time-Series

Dynamic modelling of nonlinear time-varying systems is achievable using conventional state-space analysis. For example, the state and output equations of a time-varying system can be defined as

$$\frac{d\boldsymbol{q}(t)}{dt} = \boldsymbol{A}(t)\boldsymbol{q}(t) + \boldsymbol{B}(t)\boldsymbol{x}(t) \tag{19}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}(t)\boldsymbol{q}(t) + \boldsymbol{D}(t)\boldsymbol{x}(t). \tag{20}$$

However, the complexity of devising a model of N dimensional space with time-varying coefficients, as shown in (19) and (20) as $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{C}(t)$, and $\boldsymbol{D}(t)$ is compounded by the requirement of *a priori* knowledge of each function with respect to time. Over a wide FOV such considerations must also apply as the spatially variant PSF is affected by the temporal decorrelation of the atmosphere. Adopting Taylor's frozen turbulence hypothesis [14] does not address the issue of turbulence evolution over a wide aperture. The adaptation of a RNN was employed in an attempt to capture this quality for prediction of the SVPSF. A discussion on the methods used to verify this are provided in Section VI-D.

VI. SIMULATIONS

Simulations were conducted in two phases. The first phase required the propagation of 800 nm light from reference and target beacons using the phase screen method [33]. A curvature sensor extracted 19 Zernike coefficients from two intensity pattern images, each pair comprising a data ensemble of 1000 frames. The second phase employed several hundred data ensembles, firstly for the training, and finally for testing. The MSE was then recorded, based on the prediction error between the actual and estimated aberrations, as the separations between the target object and reference beacons were varied over an anisoplanatic region. The following subsections detail the methods used in both phases.

A. Overview

The image plane can be represented by three isoplanatic regions and one anisoplanatic region over a 2 arc minute, 20 arc second FOV, as shown in Fig. 5. A boundary atmospheric layer is represented by a phase screen and perturbs target objects and source (NGS) beacons, $T_{1...n}$ and $S_{1...n}$ respectively, in the image plane. Each source beacon has a corresponding target object within an isoplanatic region Δ, Ξ and Λ; each region is defined by angular separations of $0 \leq \theta_0 \leq 20$ arc seconds. An anisoplanatic region Υ supports extended target objects, $E_{1...n}$. Within each isoplanatic region, perturbations acquired from a source beacon can be used to effectively compensate turbulence affecting a corresponding target object within an isoplanatic patch. However, none of the source beacons in any of the isoplanatic regions can be used to individually estimate
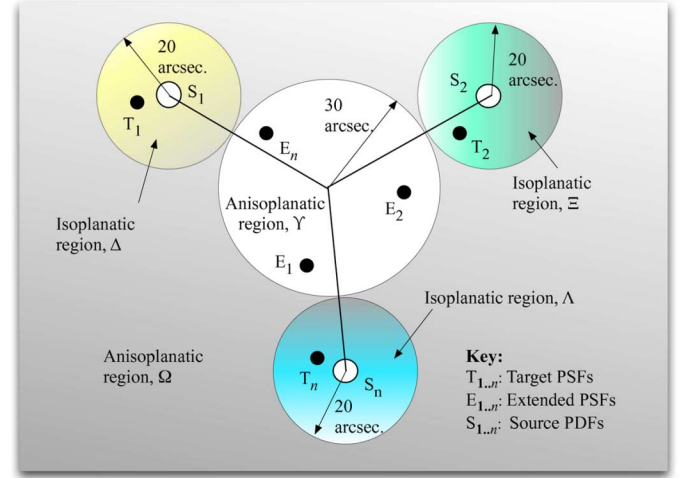

Fig. 5. Simulation platform representing a multi-aperture image plane over a wide field-of-view.

any of the extended target objects within the anisoplanatic region, Υ. Used collectively however, all N source objects were required to estimate aberrations affecting any of the extended target objects within region Υ.

Time-series data that resulted from several simulation runs supported the analysis presented in Section V. A detailed description of our simulation platform, including phase screen generation and estimation of wavefront aberrations, is provided in Section VI-B.

Simulating the propagation of N source objects through multiple turbulent layers and recovering the Zernike coefficients a_i given by (12), where $i = 2 \ldots 20$, forms the basis for our work to predict the spatially-variant PSF over a wide FOV.

B. Propagation Through Turbulence

A phase screen was required to provide Kolmogorov power spectrum of refractive index fluctuations over a wide FOV in excess of 3 arc minutes. The random midpoint displacement method [33] was used to generate a phase screen, based on a telescope diameter, D, of 1 m and the Fried coherence length, r_0, of 0.05. Thus, $D/r_0 = 20$. The resulting phase screen is shown in Fig. 6(a).

Regions representing individual, isoplanatic *patches* were selected over the phase screen and the propagation of monochromatic light from a NGS was used to create independent wavefront aberrations in the pupil plane. The extraction of two isoplanatic *patches* is shown in Fig. 6(b). The curvature method by Roddier [3] was used to recover the wavefront aberrations in the image plane and the Zernike terms, $Z_{2...20}$, for the NGS reference beacon shown in the top-left of Fig. 6(b), is shown in Fig. 6(d).

A time series ensemble comprising atmospherically distorted, stationary wave fronts was generated using the methods outlined in Section II-C. The phase screen was then moved with velocity, v, initially assuming the Taylor hypothesis [4]. The index of refraction fluctuations were displaced with velocity v over a fixed spatial region defined by the telescope aperture. Poisson and
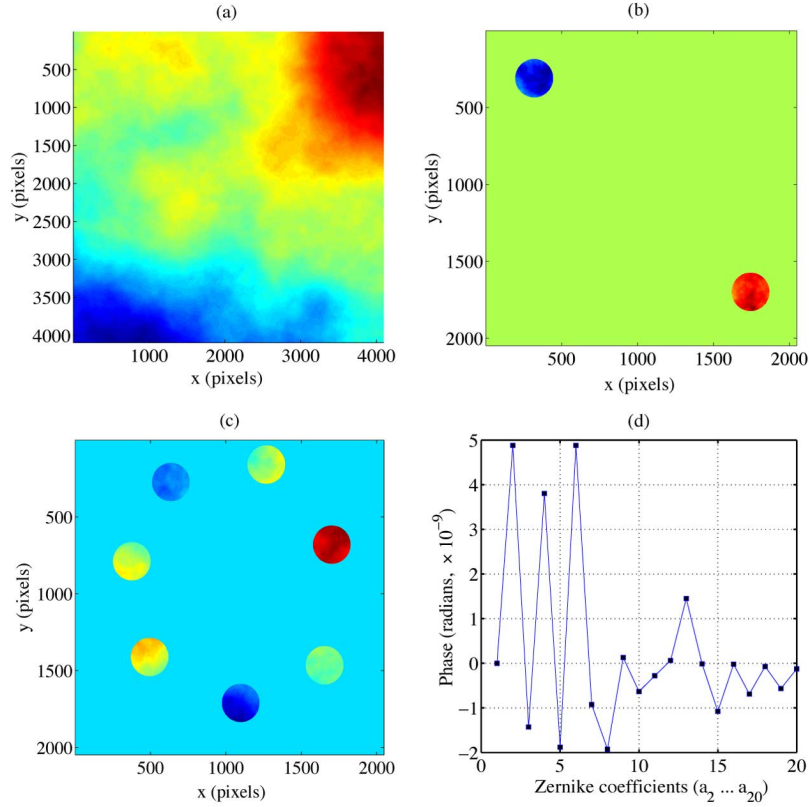
Fig. 6. (a) Phase screen, (b) multi-aperture mask: two reference beacons, (c) multi-aperture mask: seven reference beacons, and (d) Zernike coefficients $a_2 \ldots a_{20}$ from one reference beacon.

Gaussian noise models were used in recovering each isoplanatic image within each frame.

Our interest is the temporal evolution of the wavefront aberrations for real-time image restoration. Thus, later models used temporal evolution of the phase-screen, rather than a static model moved across the aperture. This resulted in the consideration of recurrent neural networks (RNNs) to provide both short and long term memory for predicting the evolution of an aberrated wave front. By incorporating both spatial and temporal data, and effectively training an ESN to learn the evolution of modal aberrations, the prediction of wavefront aberrations for PSF reconstruction in anisoplanatic regions showed to be an effective method with consistent results, in terms of MSE, over a wide spatial field. Fig. 7 shows the geometry of three NGSs α, β, γ, a target object, ϵ, and the angular separation between the target object and each source objects, $\theta_{\alpha\epsilon}$, $\theta_{\beta\epsilon}$, and $\theta_{\gamma\epsilon}$, respectively. An observational platform G, and two turbulent layers, L_2 and L_1 at altitudes z_2 and z_1 respectively, are shown with corresponding structure constants of the index of refraction fluctuations, $C_n^2(z_2)$ and $C_n^2(z_1)$.

Two objects imaged through turbulence with angular separation less that the isoplanatic angle, θ_0, are considered equivalent [4]. The isoplanatic angle, θ_0, is given by Fried [6] as,

$$\theta_0 = 58.1 \times 10^{-3} \lambda^{6/5} \left[\int_0^L C_n^2(z) z^{5/3} \, \mathrm{dz} \right]^{-3/5} \quad (21)$$
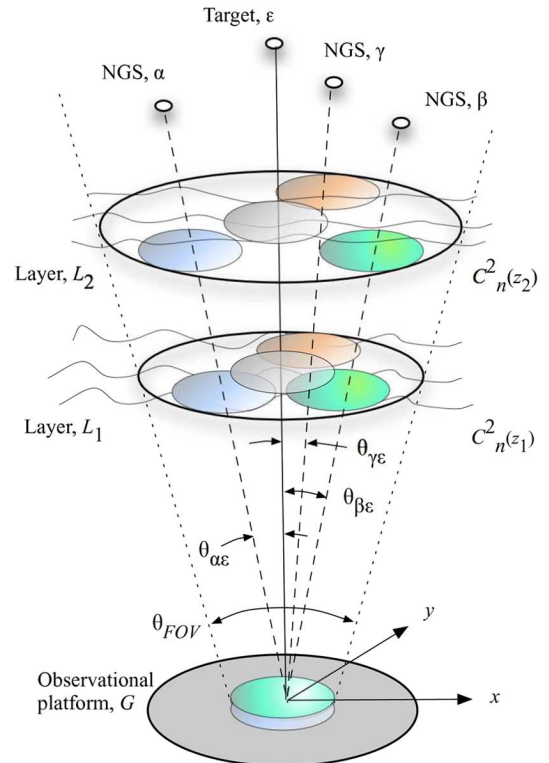

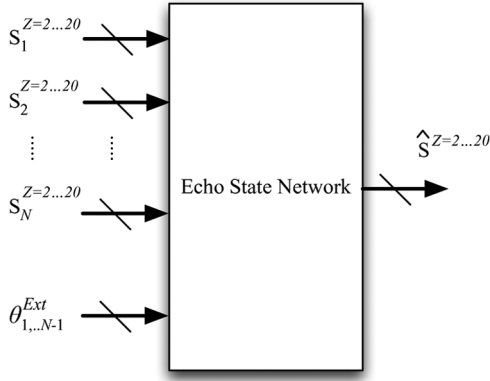Fig. 7. Geometrical view of the simulation model showing both isoplanatic and anisoplanatic regions.

Fig. 8. Echo State Network configuration used for the prediction of wavefront aberrations.

where λ is the optical wavelength, z is the altitude, C_n^2 is the structure constant of the turbulence, and L is the path length through turbulence.

The angular separation $\theta_{\gamma\epsilon}$ shown in Fig. 7 is considered to be within an isoplanatic region. However, considering the angular separations of source objects, α, and β, to target object, γ, are considered anisoplanatic. Aberrations to the PSFs that result from imaging these objects will be significantly different if either source object, α or β, was used as a reference source. This is due to imaging through an anisoplanatic region.

C. Prediction Using Reservoir Computing

A time-series ensemble, acquired from observation and simulation runs, comprised a data ensemble of 19 Zernike coefficients $Z_{2...20}$ from N source objects, $S_{1...N}$, and angular separation data, $\theta_{1...N}^{Ext}$, between a target object and each source. Each ensemble was used as a basis for time-series prediction using an echo state network (ESN). The data structure used to predict wavefront aberrations in anisoplanatic regions is shown in Fig. 8. The results of our simulations are presented and discussed in Section VII.

D. Reservoir Adaptation

The spectral radius (SR) is used as a parameter to establish the structure of the each recurrent network. The SR is the maximum eigenvalue for the dynamic reservoir's weight matrix, $\mathbf{w}$, and is not allowed to exceed unity. As there are multiple solutions to satisfy this condition, the performance of the network will vary for each reservoir structure formulated [34].

The parameters typically used for network optimization include the following.
  1) Size of the dynamic reservoir.
  2) Density of the processing elements (PEs) within the dynamic reservoir.
  3) Spectral radius.
Optimization of the ESN described in Section IV-A was considered; however, adaptations using only the spectral radius was performed. For a comprehensive discussion on reservoir adaptation see Steil [35].

A genetic algorithm has been applied to a feedforward neural network to minimize the prediction error of tilt aberrations [36]. However, due to their iterative-based architecture, we considered the gains in minimizing the prediction error would be offset by the additional computational complexity in using a generic algorithm.

VII. RESULTS

Simulations were conducted using the platform described in Section VI and the resulting data ensembles were normalized and used for training an ESN discussed in Section IV-A. From (11), Zernike coefficients a_i, where $i = 2 \ldots 20$, from N source objects, were estimated over 5 arc second incremental separations from K randomized targets within the anisoplanatic region Υ, as shown in Fig. 5. The ESN was then tested using separations from N sources within Υ that were not previously used for training. A second phase screen that provided a variable r_0 parameter was generated for this procedure. Due to the dominant effect of low-order aberrations on wavefront phase, focus (Z_4) and astigmatism (Z_5) aberrations are presented in this comparison. Separations ranging from zero to 20 arc seconds showed consistent results using multiple sources. As the isoplanatic angle was increased between 20 to 40 arc seconds within the anisoplanatic region, the average MSE remained consistent.

In order to determine the performance of ANN architecture on the spatial, as distinct to spatio–temporal effects of the aberrations, a feedforward network was included in this comparison. The effect of nonlinearity was also investigated. The response from a linear predictor on aberration data was therefore incorporated in our analysis.

Fig. 9 presents two sets of results for the prediction of defocus and astigmatism aberrations, Zernike terms Z_4 and Z_5 respectively. Fig. 9(a) and (c) show consistent prediction errors over a wide FOV using two reference beacons for Zernike terms, Z_4 and Z_5, respectively. To improve clarity only error bars for the ESN have been included on each graph. Given the selection of spectral radius, there are many possible weight matrices that comprise the same spectral radius. However, the performance, in terms of the MSE, was found to be statistically repeatable over relatively small data ensembles.

Lastly, we hypothesized that if the number of source objects were increased, the predictability of wavefront aberrations would improve. This was confirmed for seven source objects as shown in Fig. 9(b) and (d) for Zernike terms Z_4 and Z_5, respectively. However, the improvement in MSE was significantly better for the ESN, compared to the feedforward or linear predictor.

Lastly, we admit that imaging several, bright, natural source beacons within a FOV of 1–2 arc minutes is implausible. However, supplementing two or more NGSs with several artificial beacons may extend the method. Maintaining sufficient signal to noise ratio for acquisition of high order aberrations using NGSs of magnitude greater than 8, and employing the curvature sensor method as described in this article, also requires careful consideration. An example of the attenuation in phase measurements as a function of modal order is shown in Fig. 6(d).
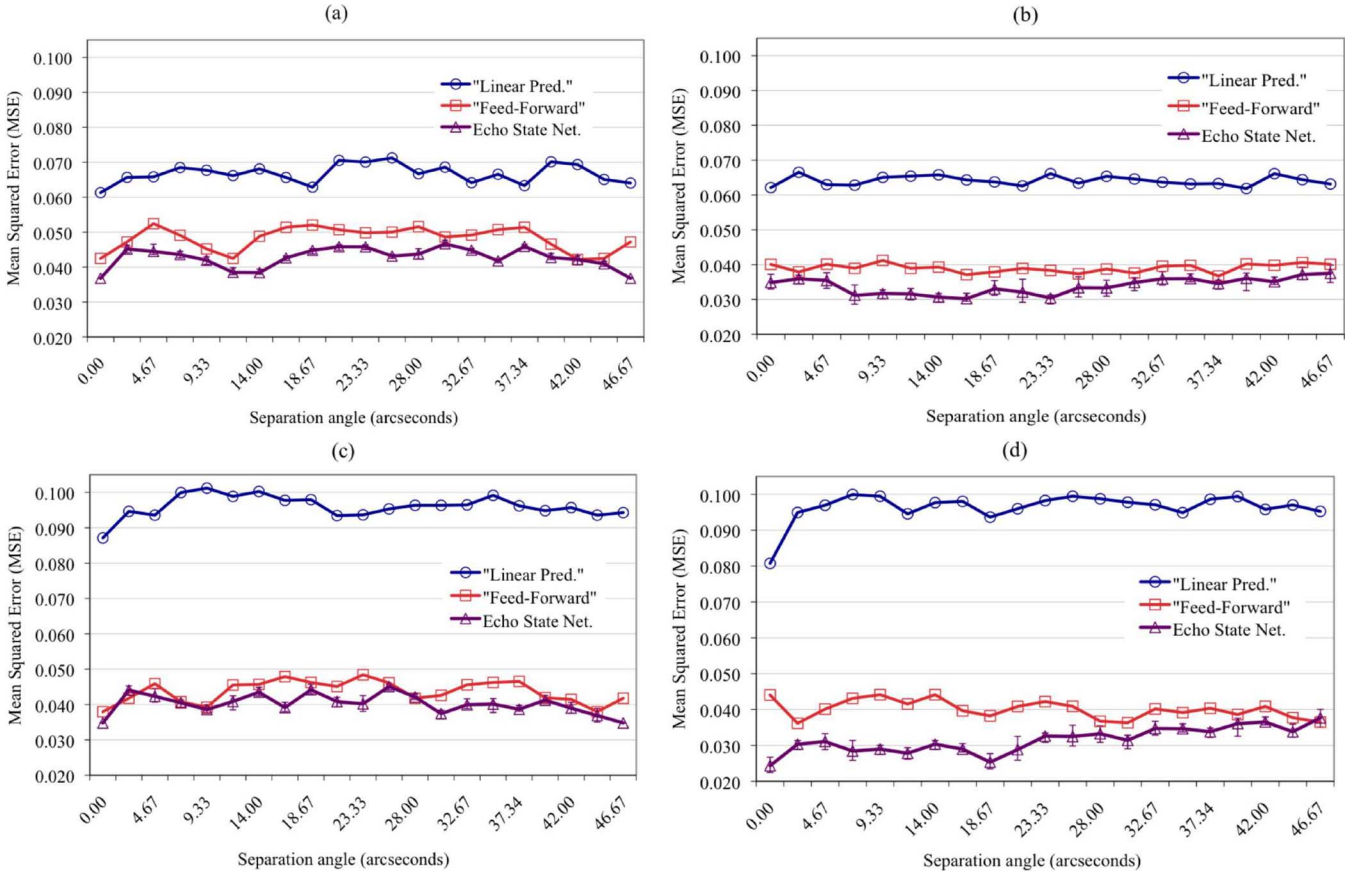
Fig. 9. MSE prediction of defocus (Z_4) and astigmatism (Z_5) Zernike terms over anisoplanatic regions, using a linear predictor, feed-forward, and echo state network: (a) Z_4 with two source objects, (b) Z_4 with seven source objects, (c) Z_5 with two source objects, and (d) Z_5 with seven source objects.

VIII. PRACTICAL CONSIDERATIONS FOR REAL-TIME APPLICATION

For effective real time application of an uncompensated DWFS system over a wide FOV, the efficiency of the method outlined in Section III for prediction of the spatially variant PSF, is dependent on the following.

- A high-speed image acquisition system supporting either a multi-aperture or wide-field configuration.
- An efficient feature extraction method that can be implemented in a field programmable gate array to extract Zernike terms from N point source objects, based on a similar method discussed in Section II-C.
- A highly parallel hardware implementation of our method outlined in Section III.

Replication of such systems using a VLSI hardware implementation for N bright point source objects can provide input data for the method described in this paper, and outlined in Fig. 8. Since ESNs only require training of their output nodes, an efficient batch training system could be achieved using *off-line* training data runs, prior to an observation session. Once trained, an optimized ESN provides a very efficient platform for predicting the spatio–temporal evolution of the PSF over a wide field. Extensions, both in terms of efficiency and system optimization through architectural enhancements, [19], [34], [37] ensure that this relatively new RNN, referred to in the literature as *reservoir computing*, will find new and emerging applications in astronomical image reconstruction and AO control systems.

IX. CONCLUSION

This paper has shown how reservoir computing can be incorporated with a spatio–temporal image model and used to predict wavefront aberrations affecting target objects within anisoplanatic regions, using wavefront data from several bright reference beacons and distributed over a wide FOV. Once these aberrations have been determined, a deconvolution algorithm can be used to reconstruct the spatially-variant PSF for the restoration of astronomical images. Since the sparse matrix of an ESN does not require training, the Moore–Penrose pseudoinverse can be used to determine output weights. A possible extension to this work is the incorporation of ANNs for classification of Zernike terms. Extensions to existing work on dimensionally reduction, referenced in this paper, is an active area of our ongoing research.

REFERENCES

[1] T. Fusco, J.-M. Conan, V. Michau, L. M. Mugnier, and G. Rousset, "Efficient phase estimation for large-field-of-view adaptive optics," *Opt. Lett.*, vol. 24, no. 21, pp. 1472–1474, 1999.

[2] R. Ragazzoni, E. Marchetti, and F. Rigaut, "Modal tomography for adaptive optics," *Astron. Astrophys*, vol. 342, pp. L53–L56, 1999.

[3] F. Roddier, *Adaptive Optics in Astronomy*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[4] M. Roggemann and B. Welsh, *Imaging Through Turbulence*. Boca Raton, FL: CRC Press, 1996.

[5] R. Hufnagel, "Optical propagation through turbulence," *Proc. Topical Mtg. on Optics*, pp. Wa1–Wa4, 1974.

[6] D. Fried, "Anisoplanatism in adaptive optics," *J. Opt. Soc. Amer.*, vol. 72, pp. 52–61, 1982.

[7] A. Quirrenbach, "Adaptive optics with laser guide stars: Basic concepts and limitations," in *Laser Guide Star Adaptive Optics for Astronomy*, N. Ageorges and C. Dainty, Eds. Dordrecht, The Netherlands: Kluwer, 1997, vol. 551, pp. 23–50.

[8] M. Tallon and R. Foy, "Adaptive telescope with laser probe—isoplanatism and cone effect," *Astron. Astrophys.*, vol. 235, pp. 549–557, Aug. 1990.

[9] R. R. Ragazzoni, E. Marchetti, and G. Valento, "Adaptive-optics corrections available for the whole sky," *Nature (UK)*, vol. 403, no. 6765, pp. 54–56, Jun. 2000.

[10] M. A. v. Dam and R. G. Lane, "Extended analysis of curvature sensing," *J. Opt. Soc. Amer.*, vol. 19, no. 7, pp. 1390–1397, 2002.

[11] P. McGuire, D. Sandler, M. Lloyd-Hart, and T. Rhoadarmer, "Adaptive optics: Neural network wavefront sensing, reconstruction, and prediction," in *Scientific Applications of Neural Nets, Proceedings of the 194th W. E. Heracus Seminar, 1998*, J.W. Clark, T. Lindenau, and M. L. Ristig, Eds. Berlin, Germany: Springer-Verlag Publishers, 1999, pp. 97–97.

[12] G. Chundi, M. Lloyd-Hart, and M. Sundareshan, "Training multilayer perceptron and radial basis function neural networks for wavefront sensing and restoration of turbulence-degraded imagery," in *2004 IEEE Int. Joint Conf. Neural Networks*, Budapest, Hungary, 2004, vol. 3, pp. 2117–2122.

[13] S. Engelberg, *Random Signals and Noise—A Mathematical Introduction*. Boca Raton, FL: CRC, 2007.

[14] J. W. Goodman, *Introduction to Fourier Optics*. New York: McGraw-Hill, 1968.

[15] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Process. Mag.*, vol. 14, no. 2, pp. 24–41, Mar. 1997.

[16] M. Aubailly, M. C. Roggermann, and T. J. Schulz, "Prediction of the space-varying point spread function for reconstruction of anisoplanatic adaptive optics images," in *Astronomical Adaptive Optics Systems and Applications II*, R. Tyson and M. Lloyd-Hart, Eds. Bellingham, WA: SPIE—The International society for Optical Engineering, 2005, vol. 5903.

[17] T. Fusco, J.-M. Conan, L. M. Mugnier, V. Michau, and G. Rousset, "Characterization of adaptive optics point spread function for anisoplanatic imaging. Application to stellar field deconvolution," *Astron. Astrophys.*, vol. 142, pp. 149–156, Feb. 2000.

[18] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *Opt. Soc. Amer.*, vol. 66, pp. 207–211, 1976.

[19] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, pp. 78–80, 2001.

[20] D. P. Mandic and J. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. New York: Wiley, 2001.

[21] S. J. Weddell and R. Y. Webb, M. Cree, Ed., "A neural network architecture for the reconstruction of turbulence degraded point spread functions," in *Proc. Image & Vision Computing NZ 2007 (IVCNZ'07)*, Hamilton, Dec. 2007, pp. 103–108.

[22] M. Lukosevicius and H. Jaeger, Overview of Reservoir Recipes School of Engineering and Science, Jacobs Univ. Bremen, Germany, Tech. Rep. 11, 2007 [Online]. Available: http://www.jacobs-university.de/imperia/md/content/groups/research/techreports/reservoiroverview_techreport11.pdf

[23] S. J. Weddell and R. Y. Webb, I. Press, Ed., "Dynamic artificial neural networks for centroid prediction in astronomy," in *Sixth Int. Conf., Hybrid Intelligent Systems (HIS'06)*, Auckland, New Zealand, Dec. 2006, pp. 68–72.

[24] D. McGaughey and G. M. Aitken, "Temporal analysis of stellar wavefront-tilt data," *J. Opt. Soc. Amer. A*, vol. 14, no. 8, pp. 1967–1974, 1997.

[25] R. K. Tyson, "Introduction to adaptive optics," *Opt. Eng.*, vol. TT41, 2000.

[26] A. Fuchs, M. Tallon, and J. Vernin, "Focusing on a turbulent layer: Principle of the generalized scidar," *Pub. Aston. Soc. Pacific (PASP)*, vol. 110, pp. 86–91, 1998.

[27] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, pp. 43–50, 1997.

[28] U. Parlitz, I. Wedekind, W. Lauterborn, and C. Merkwirth, TSTOOL and User Manual, ver. 1.11 Univ. Göttingen, 2002 [Online]. Available: http://www.physik3.gwdg.de/tstool/index.html

[29] A. Wolf, J. Swift, L. Swinney, and J. Vastano, "Determining lyapunov exponent from a time series," *Physica D*, vol. 16, pp. 285–317, 1985.

[30] J. C. Sprott, *Chaos and Time-Series Analysis*. Oxford, U.K.: Oxford Univ. Press, 2003.

[31] D. A. Montera, B. M. Welsh, M. C. Roggemann, and D. W. Ruck, "Prediction of wave-front sensor slope measurements with artificial neural networks," *Appl. Opt*, vol. 36, no. 3, pp. 675–681, 1997.

[32] M. Lloyd-Hart and P. McGuire, "Spatio–temporal prediction for adaptive optics wavefront reconstructors," *Proc. ESO*, pp. 95–101, 1995.

[33] C. M. Harding, R. A. Johnston, and R. G. Lane, "Fast simulation of a kolmogorov phase screen," *App. Opt*, vol. 38, no. 11, pp. 2161–2170, 1999.

[34] M. Ozturk, D. Xu, and J. Principe, "Analysis and design of echo state networks," *Neural Comput.*, vol. 19, pp. 111–138, 2007.

[35] J. J. Steil, "2007 special issue: Online reservoir adaptation by intrinsic plasticity for backpropagation-decorrelation and echo state learning," *Neural Netw.*, vol. 20, no. 3, pp. 353–364, 2007.

[36] P. Gallant and J. Aitken, "Genetic algorithm design of complexity-controlled time-series predictors," in *IEEE XIII Workshop on Neural Networks for Signal Processing*, 2003, pp. 769–778.

[37] R. Webb, "A cascade architecture for on-line predictive echo state networks," *MLMTA*, pp. 99–104, 2007.

Stephen Weddell (M'06) received the B.Sc. degree from Curtin University, Perth, Western Australia, in 1991 and the M.Sc.(Eng.) degree from the same university in 1997. He is currently working towards the Ph.D. in electrical and electronic Engineering at the University of Canterbury, Christchurch, New Zealand.

Since 2001, he has been a Lecturer on computer hardware and applied digital signal processing in the Department of Electrical and Computer Engineering, University of Canterbury.

Russell Y. Webb (M'05) received the Ph.D. in electrical engineering from Cornell University, Ithaca, NY, in 2000.

He was a Software Engineer working on operating system development, security, and advanced technology before joining the faculty of Electrical and Computer Engineering at the University of Canterbury, Christchurch, New Zealand, in 2004. Currently, he is conducting research in predictive machine learning.