

Reconstructing Velocities of Migrating Birds from Weather Radar – A Case Study in Computational Sustainability

The Faculty of Oregon State University has made this article openly available.
Please share how this access benefits you. Your story matters.

Citation	Farnsworth, A., Sheldon, D., Geevarghese, J., Irvine, J., Van Doren, B., Webb, K., . . . Kelling, S. (2014). Reconstructing velocities of migrating birds from weather radar - A case study in computational sustainability. <i>AI Magazine</i> , 35(2), 31-48. doi:10.1609/aimag.v35i2.2527
DOI	10.1609/aimag.v35i2.2527
Publisher	Association for the Advancement of Artificial Intelligence (AAAI)
Version	Accepted Manuscript
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Reconstructing Velocities of Migrating Birds from Weather Radar – A Case Study in Computational Sustainability

**Andrew Farnsworth¹, Daniel Sheldon²,
Jeffrey Geevarghese², Jed Irvine³, Benjamin Van Doren¹, Kevin Webb¹,
Thomas G. Dietterich³ Steve Kelling¹**

¹Cornell Lab of Ornithology, Ithaca, NY, USA; {af27,bmv25,kfw4,stk2}@cornell.edu

²University of Massachusetts Amherst, Amherst, MA, USA; {sheldon,jeffrey}@cs.umass.edu

³Oregon State University, Corvallis, OR, USA; {irvine,tgd}@eecs.oregonstate.edu

Bird migration occurs at the largest of global scales, but monitoring such movements can be challenging. In the US there is an operational network of weather radars providing freely accessible data for monitoring meteorological phenomena in the atmosphere. Individual radars are sensitive enough to detect birds, and can provide insight into migratory behaviors of birds at scales that are not possible using other sensors. Archived data from the WSR-88D network of US weather radars hold valuable and detailed information about the continent-scale migratory movements of birds over the last 20 years. However, significant technical challenges must be overcome to understand this information and harness its potential for science and conservation. We describe recent work on an AI system to quantify bird migration using radar data, which is part of the larger *BirdCast* project to model and forecast bird migration at large scales using radar, weather, and citizen science data.

Introduction

Ecological processes that occur at global scales are among the most impressive natural phenomena on Earth; they link environments across the globe and capture the imagination of observers. For example, billions of North American birds, representing more than 400 species, migrate annually, often thousands of miles, frequently under the cover of darkness, between breeding and non-breeding areas. There is an urgent need to understand global processes in order to mitigate threats from anthropogenic change. For example, man-made collision hazards are killing hundreds of millions of birds annually (Crawford et al. 2001; Longcore et al. 2008), and direct and indirect habitat alteration is reducing stopover habitats vital for birds to refuel during migration (Bonter et al. 2009; Veltri and Klem 2005). Our ability to mitigate this anthropogenic mortality relies on a detailed understanding of the magnitude and composition of migration in time and space so we can take actions such as protecting critical stopover sites and turning off wind-turbines and lights on man-made structures at key places and times during migration.

However, because of their scale and complexity, global phenomena are also very difficult to study. Bird migration is one of the most complex and dynamic natural systems on the planet and it presents immense challenges to observation and measurement. Despite extensive research (e.g., for reviews see Deppe and Rotenberry 2008, and Moore et al., 2005) we do not have the necessary information to accurately model and predict continent-wide migration. Our knowledge of migration is largely based on studies from specific locations (Able 1999; Veltri and Klem 2008; Packett and Dunning 2009) or studies that mark or track individual birds with leg-bands or tracking devices (Francis 1995; Schaub et al. 2004; Tøttrup et al. 2006; Newton 2008). Despite technological advances, these approaches are costly and provide detailed movement information for only a small number of birds. New data sources and new approaches are needed to understand migration at the largest of scales.

A Way Forward: Computational Sustainability

Three approaches in the new field of Computational Sustainability enable the study global ecological phenomena: improved sensors, citizen science, and repurposing existing sources. Each of these approaches requires advances in AI, automated reasoning, and machine learning.

The first new approach is improved sensors. Computer scientists and engineers, in collaboration with domain scientists, are developing novel sensors and sensor platforms that can collect much more data than in the past. Examples include deep ocean gliders (Eriksen et al. 2001; Sherman et al. 2001) that can collect data throughout the oceans, autonomous aerial vehicles that can collect high-resolution data on ecosystems (Carey 2012), and environmental sensor networks (NEON 2012). The deployment of novel sensors (particularly via autonomous robots) requires advances in robot navigation and decision-making, as well as the ability to fit scientific models to the resulting data. Unfortunately, for bird migration, existing sensors that are small enough and light enough to be carried by birds can be placed on only a tiny fraction of the billions of individual birds that are migrating.

The second approach is citizen science. For example, the eBird project is a web-based platform that collects observations from volunteer bird watchers and presently archives over 140,000,000 observations

from every country in the world. By applying novel machine learning techniques to these data, our collaborators have been able to construct high-resolution maps of bird habitat throughout the United States and broader parts of the Americas (Fink et al. 2010; Fink et al. 2013). Analyzing citizen science data requires advances to overcome challenges of data quality and unevenly-

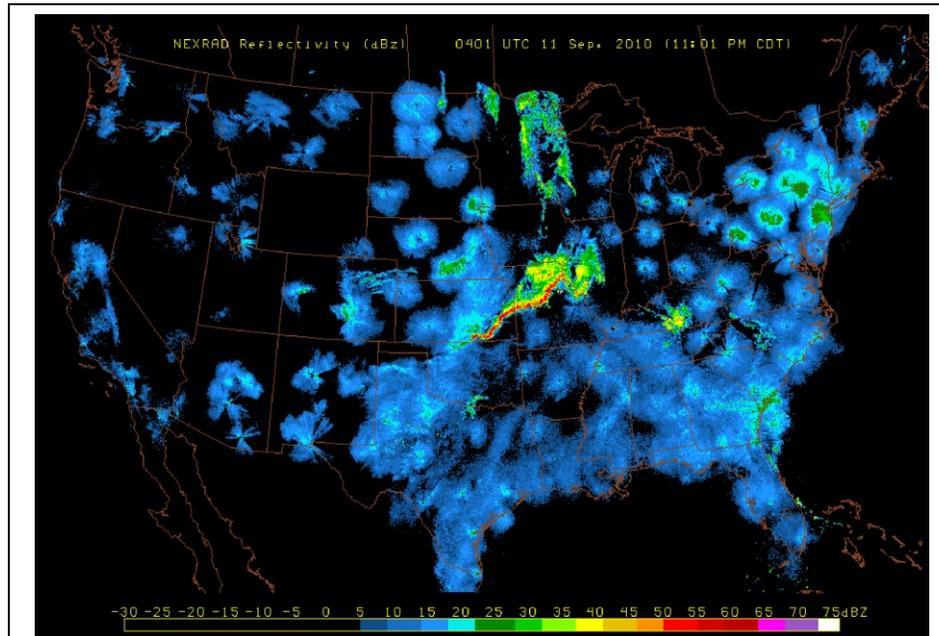


Figure 1: Mosaic image of the continental US network of WSR-88D stations depicting a heavy migration event on night of 10 September 2010. Precipitation appears as contiguous regions with blocky and irregular shapes. Birds appear as more uniform stippled patterns of primarily light and medium grays surrounding individual radar stations. Bird migration is evident across the entire continent, particularly across the central and eastern US. Darker stippled areas represent bird densities approaching 2,000 birds km³. On this night, many tens of millions of birds were aloft over the US. Note, also, that areas with the uniform stippling do not necessarily represent areas where no birds are present, rather these are places either without radar coverage or where the radar beam is sufficiently far and high from the radar so as not to provide useful information

distributed data, since volunteers may have vastly different levels of expertise (Kelling et al. 2012) and they choose when and where to make observations.

The third approach is to repurpose existing data sources so that they can be applied to address ecological problems. The work described in this paper is a case study of this approach. Below, we describe the challenges of analyzing weather radar data to extract information on bird density and migration velocity.

Our Case Study: Reconstructing Velocities of Migrating Birds

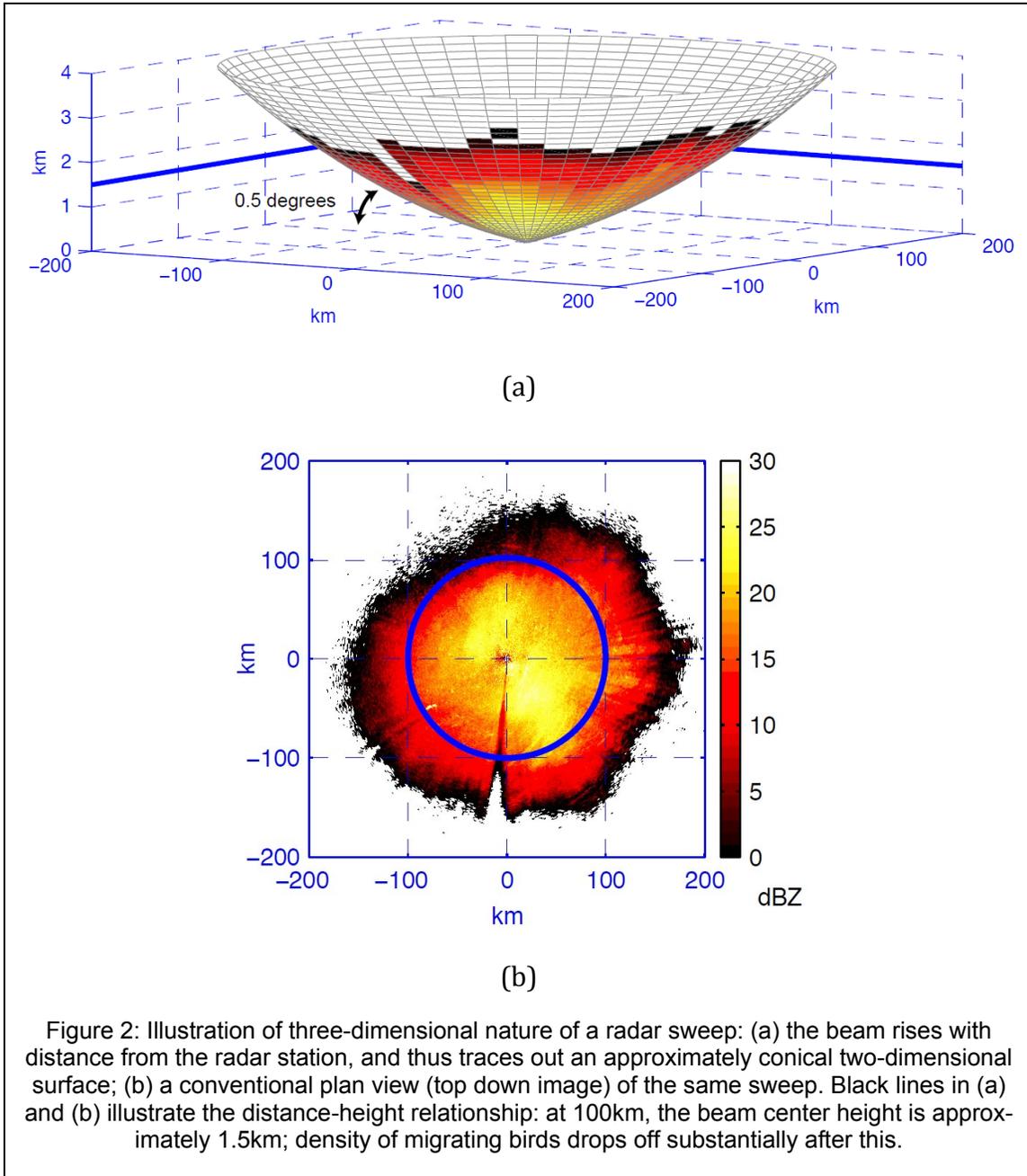
The National Weather Service operates the WSR-88D (Weather Surveillance Radar-1988 Doppler) network of 159 radar stations in the United States and its territories. The network covers nearly the entire US and was designed to detect and study weather phenomena such as precipitation and severe storms (Crum and Alberty 1993; Doviak and Zrnic 1993). However, it turns out that radar systems also detect birds, bats, and insects in the atmosphere (Lack and Varley 1945), which makes WSR-88D an ideal source for information about movements of birds (Figure 1). Consequently, a small community of researchers has been pursuing the idea of using weather radar as a tool for observing the presence and movement of migrating birds (Gauthreaux and Belser 2003; Dokter et al. 2011). Fortunately, the WSR-88D data are archived and available from the early 1990s to present. However, despite the fact that several studies have examined WSR-88D data (e.g. Gauthreaux, Belser, and Van Blaricom 2003; Buler et al. 2012), these analyses were limited in scope, either to a small region of space (e.g. a single or a few station/s) or to short time periods (e.g. a few weeks out of a season). If it is possible to develop the necessary algorithms to fully automate radar data analysis, then this lengthy time series could enable us to study bird migration over a period of more than 20 years, which could be invaluable for building a scientific understanding of bird migration.

A key challenge limiting the scope of previous studies is that WSR-88D data are difficult to interpret. Radar signatures of birds are quite similar to those of bats, insects, meteorological phenomena (e.g., precipitation), and even airborne dust and pollen. Consequently, a highly trained expert must review and interpret these data (Gauthreaux et al. 2003; Buler and Diehl 2009). This is not feasible unless the data used are significantly restricted in scope, either spatially or temporally. Assuming a single radar scans its surrounding atmosphere every five to ten minutes, we estimate there are well over 100 million archived volume scans from single radar sites, and a single night during the peak of bird migration will produce about 15,000 scans nationwide. If we can develop AI tools to automate the interpretation steps, this vast store of data could enable important advances in our understanding of many phenomena, particularly bird migration. For example, how many birds migrate on a given night? How does the total density of migrants change over time across a migration season or across years, or even across the continent during the last two decades? In what directions are birds moving and at what speed? With what local and regional conditions do large movements correlate?

In this article, we describe ongoing work to develop an AI system that automatically interprets radar data and extracts quantitative high-level information about migration. A primary goal of the article is to provide an accessible overview of a recent technical paper that introduced a new probabilistic model and approximate inference algorithm for reconstructing the velocities of migrating birds from the partial information collected by Doppler radar (Sheldon et al. 2013). An accurate algorithm to estimate the velocities of birds and other targets detected by radar is critical to unlocking the potential of the data. Velocity information is important for understanding the biology of bird migration. In addition, there is growing evidence that the structure of the velocity field is a key to discriminating between birds, which fly under their own power, and other targets such as

precipitation, insects, and dust, that are primarily carried by the wind (Gauthreaux et al. 2003; Dokter et al. 2011).

The article also highlights key themes that relate to Computational Sustainability and AI. One theme is the promise of data mining and machine learning techniques for interpreting large eco-



logical data sets such as those collected by sensor networks or citizen scientists (Kelling et al. 2012). Because the availability of large data sets in ecology is relatively recent, they hold great promise for scientific discovery, and their analysis raises unique and unsolved technical challenges (Kelling et al. 2012; Dietterich et al. 2012). Close collaborations between domain scientists and computer scientists are a fertile environment for innovation. Another theme is the power of joint

probabilistic inference over a purely sequential approach in AI systems that consist of a pipeline of processing steps, which is discussed in Section 4 after providing background on the velocity inference problem and existing models from the weather community.

Radar Background and Problem Statement

Scanning Strategy

Each radar in the WSR-88D network collects data by conducting a sequence of *volume scans*, which complete every six to ten minutes. Each volume scan consists of a sequence of *sweeps* during which the antenna rotates 360 degrees around a vertical axis while keeping its elevation angle fixed (Figure 2). The result of each sweep is a set of raster data products summarizing the radar signal returned from targets within discrete *pulse volumes*, which are the portions of the atmosphere sensed at a particular antenna position and range from the radar. The coordinates (r, ϕ, ρ) of each pulse volume are measured in a three-dimensional polar coordinate system:

- r is the distance in meters from the antenna,
- ϕ is the *azimuth*, which is the angle in the horizontal plane between the antenna direction and a fixed reference direction (typically degrees clockwise from due north),
- ρ is the elevation angle, which is the angle between the antenna direction and its projection onto the horizontal plane.

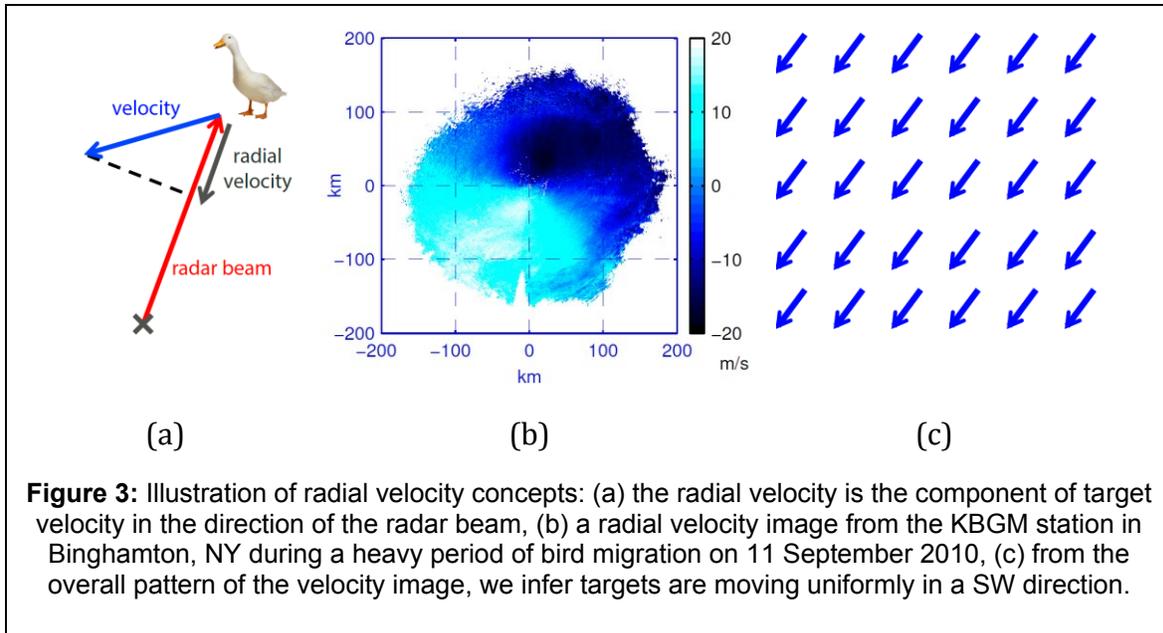
Figure 2(a) illustrates an important aspect of the radar’s volume scanning strategy. The upward angle of the beam and the earth’s curvature cause the beam to sample higher points in the atmosphere as it travels away from the radar, which results in a distance–height relationship that is important to remember when interpreting conventional “top-down” radar images like the one in Figure 2 (b). In this example, the beam encounters migrating birds whose density drops off sharply above altitudes of about 1.5 km; thus, the radar can only “see” the birds out to distances of about 150 km, after which the beam is too high. This geometry explains the appearance of the bird echoes in the national composite radar map of Figure 1. Each circular region represents the birds detectable from a single radar station, leading to an overall picture of migration activity that appears localized and uneven. However, a more accurate interpretation is that a fairly even blanket of migrating birds covers much of the US.

For this article, we will illustrate radar concepts using images from the lowest sweep ($\rho = 0.5$ degrees), as shown in Figure 2. However, the radar conducts sweeps at a sequence of increasing elevations, and thus samples the three-dimensional volume surrounding the station using concentric cones like this one in Figure 2 (a). Our algorithms use data from all sweeps.

Radial Velocity

The two primary data products collected from each sweep are *reflectivity*, a measure of the total amount of power returned to the radar from targets within each pulse volume, and *radial velocity*, the average speed at which the targets in the pulse volume are approaching or departing the radar, which is measured using the Doppler shift of the radar signal. While reflectivity is the measurement most familiar from weather reports, because it indicates the amount of precipitation (or birds) in the atmosphere, we are primarily concerned with radial velocity in this article.

Figure 3 illustrates the interpretation of WSR-88D radial velocity data. For any given pulse volume, radial velocity tells us the component of target velocity in the direction of the radar beam,



and we have no additional information about the component orthogonal to the radar beam. However, the overall pattern of the sweep often provides clear evidence about the true target velocities. In this example, targets to the NE of the radar station have negative radial velocities (dark colors), which means they are approaching the radar, and targets to the SW of the radar station have positive radial velocities (light colors), which means they are departing the radar station. We can infer that the targets (in this case, predominantly migrating birds) are moving uniformly in a SW direction, as shown in panel (c). The spiral pattern in the velocity image is due to changes in target velocity with height, which are partly driven by changes in wind direction. In this case, winds at lower elevations are from the north, whereas those at higher elevations have more of an easterly component, which explains why the boundary between positive and negative radial velocities rotates clockwise with distance from the radar.

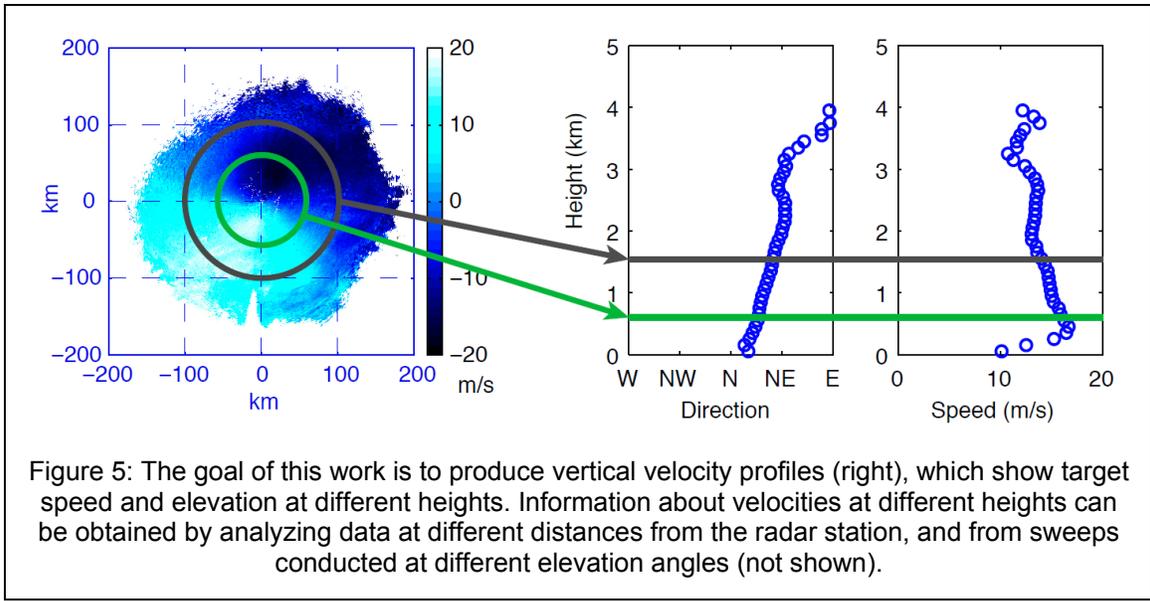
Problem Statement

Our goal is to infer the true velocities of migrating birds at different elevations from WSR-88D radial velocity data. Specifically, the final output of our algorithms will be elevational profiles like the ones shown in Figure 4 that show the migration direction and speed in each elevation bin.

In the following sections, we will introduce the basic modeling approaches from the weather community for solving this problem, and highlight a key difficulty that arises due to the phenomenon of *aliasing* in the radial velocity data. We then propose a modification to the model that is conceptually simple but leads to a much more difficult inference problem, and then we develop an inference algorithm using approximate inference techniques proposed recently in the AI and machine learning communities.

Existing Models and the Problem of Aliasing

How can one reconstruct the true direction and speed of migrating birds from the partial information contained in radial velocity data? Our interpretation in Figure 3 relied on the overall pattern of the radial velocity image combined with an implicit assumption that targets (in this case, birds) behaved in a reasonably consistent way, and thus had similar velocities, throughout the

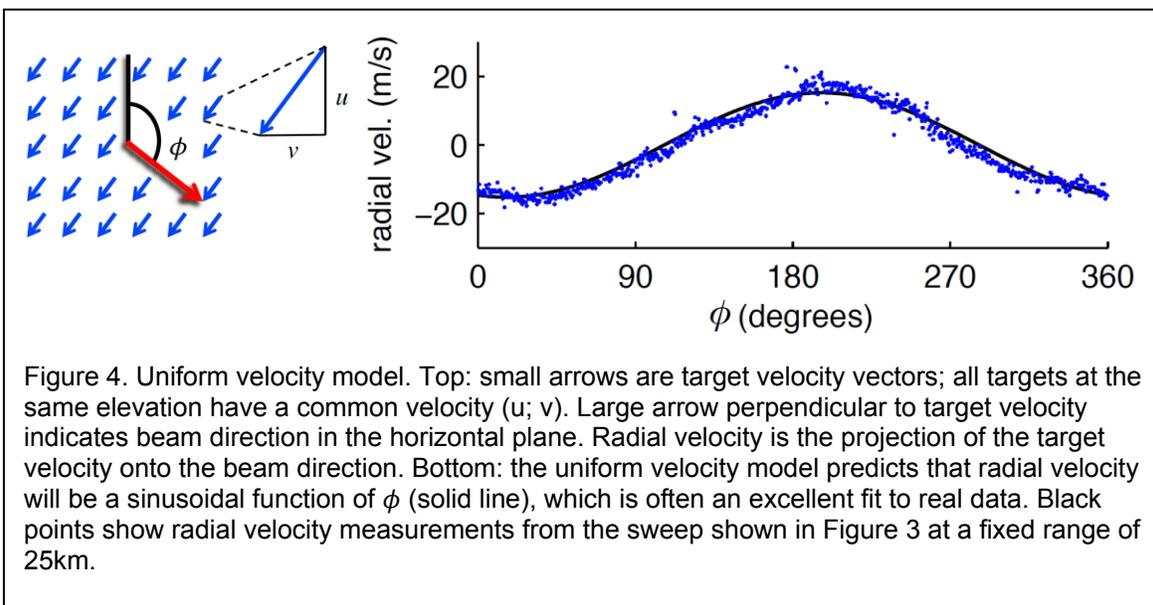


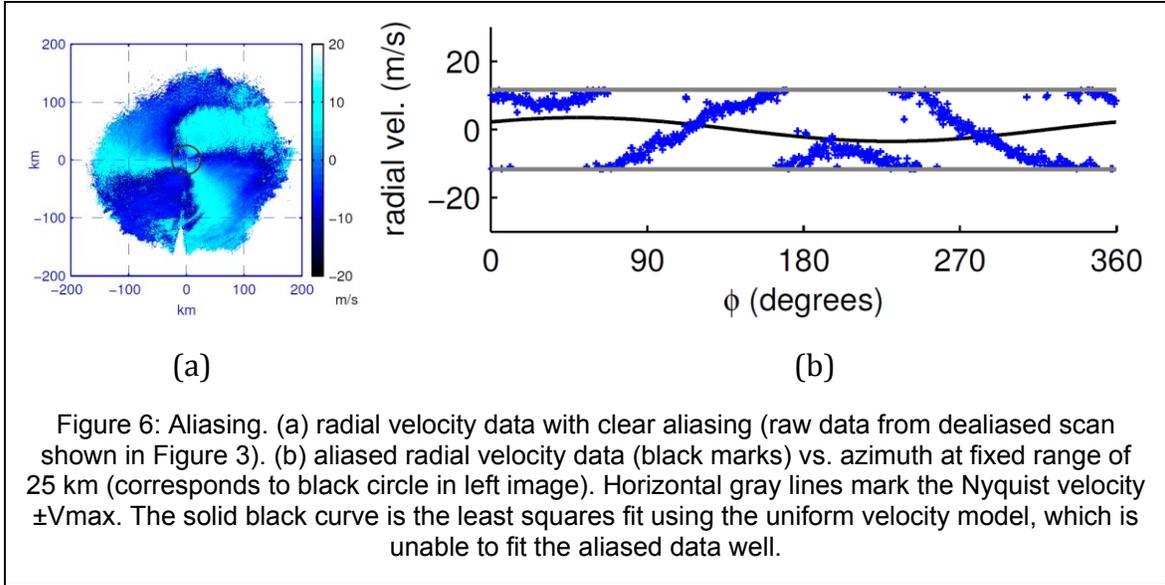
region covered by the sweep. This is a basic assumption made by radar-based wind profiling algorithms, and it will also be the foundation for our algorithms.

Velocity Model. Our starting point is the *uniform velocity model* from the weather community (Doviak and Zrnic 1993), which assumes that target velocity varies only with elevation, so it is constant within a given elevation range. Suppose the true unknown velocity components in the x and y directions of the horizontal plane are u and v .¹ Then, for a fixed elevation angle ρ , it is a simple exercise in trigonometry to determine the radial velocity at a given antenna angle ϕ in the horizontal plane. Figure 5 illustrates the situation.

The radial velocity $\mu(\phi)$ predicted by this model is the projection of the unknown velocity vector onto the beam direction, which is (Doviak and Zrnic 1993; Dokter et al. 2011):

$$\mu(\phi) = u \cos \phi \cos \rho + v \sin \phi \sin \rho.$$





Two things are notable about this model. First, it describes a sinusoidal function of ϕ .² As seen in Figure 5, real data is often an excellent fit with this model, but exhibits some noise due to fine-scale unmodeled variability in the velocity field and measurement noise. Thus, when fitting the model it is natural to model the true radial velocity $a(\phi)$ as a Gaussian random variable with mean $\mu(\phi)$:

$$a(\phi) \sim \mathcal{N}(\mu(\phi), \sigma^2).$$

The second notable feature of the model is that $\mu(\phi)$ is a *linear* function of the unknown velocity components u and v , so the maximum-likelihood parameters can be recovered by a simple least squares fit. Variants of this fitting approach are referred to either as *velocity-azimuthal display* (VAD) algorithms or *velocity volume profiling* (VVP) (Doviak and Zrnic 1993).

Aliasing. The uniform velocity model is an excellent model for our purposes and it is easy to fit. There would be little more to say if not for the problem of *aliasing*. Aliasing occurs when the target radial velocities exceed a value V_{max} called the *Nyquist velocity*. Due to the sampling strategy for measuring the Doppler shift of the returned signal, WSR-88D radars can only resolve radial velocities unambiguously if they are smaller than V_{max} . When the true radial velocity exceeds this value, the measurement will “wrap around”. For example, the Nyquist velocity is $V_{max} = 11.61 \text{ ms}^{-1}$ in one of the common clear-air operating modes that is used during conditions favorable for bird migration. For an outbound target whose true radial velocity is $V_{max} + 1 = 12.61 \text{ ms}^{-1}$, the measurement will instead read as $-V_{max} + 1 = -10.61 \text{ ms}^{-1}$, so it appears instead as a fast *inbound* target. More precisely, any radial velocity value a that exceeds the Nyquist velocity is *aliased*³ by shifting it by an integer multiple of $2V_{max}$ to a new value \bar{a} that falls in the interval $[-V_{max}, V_{max}]$, which we will call the *unambiguous interval*. Thus, an infinite number of different radial velocities can lead to the same measurement: these are called *aliases*.

²The phase and amplitude are determined by the unknown velocity parameters u and v .

³The aliasing operation of a value a is mathematically equivalent to the modulus operation $a \bmod 2V_{max}$ in modular arithmetic, except it follows the convention that the result lies in the interval $[-V_{max}, V_{max}]$ instead of $[0, 2V_{max}]$.

In fact, aliasing errors were present and had to be corrected in the velocity data for the example we interpreted in Figure 3. The raw data for the same sweep are shown in Figure 6 (top). Aliasing is clearly evident: the fastest moving inbound targets to the NE of the radar are erroneously measured as having outbound velocities (light), while the fastest outbound targets to the SW are erroneously measured as having inbound velocities (dark). Even more important to our modeling goals is the fact that the uniform velocity model completely breaks under aliasing. Figure 6 (bottom) shows the effect of aliasing on the noisy sine curve of the uniform velocity model: the curve now wraps around at $\pm V_{max}$. A least squares fit of the uniform velocity model directly to this data will give very poor results (solid curve in the figure). For this reason, most approaches to velocity profiling assume that the data has already been properly dealiased.

A Joint Model for Velocity Profiling with Aliased Data

How can we proceed in the presence of aliasing errors? While most previous approaches attempt to *fix the data* by dealiasing it before fitting the velocity parameters, it is much more desirable to *fix the model* instead, so that it can jointly reason about each source of missing information in this problem: (1) the unobserved component of radial velocity, (2) ambiguity due to aliasing, and (3) noise due to fine-scale variability in the velocity field.

The argument in favor of a joint model illustrates a general principle for AI systems that perform a sequence of processing steps where the output from one step feeds into the next: when possible, joint inference over the entire pipeline is preferable to solving each of the problems separately in a feed-forward fashion, *if* one can meet the more challenging computational demands of joint inference (Poon and Domingos 2007). If the tasks are treated separately, errors in earlier stages propagate and compound in later ones—there is no opportunity to correct dealiasing errors during velocity profiling. Joint inference not only avoids this problem, but it also leverages useful higher-level information from later stages to guide the decisions made in earlier stages. In our case, the velocity model provides excellent advice about the correct interpretation of aliased data: knowing that the data in Figure 6 should follow a noisy sine curve, it is visually obvious where the sine curve crosses an aliasing boundary and how the data should be adjusted to fix this. This is much more powerful than the weaker spatial continuity assumptions used by many existing dealiasing algorithms. Other dealiasing algorithms use separate wind measurements to provide information about the expected velocities—this can be viewed as a surrogate for jointly reasoning about aliasing and velocity, but the major downside is limited availability of separate wind measurements.

Fixing the Uniform Velocity Model

Empirically, the uniform velocity model performs very poorly when aliasing errors are present. But what goes wrong conceptually? By examining the problem, we will find a way to fix the model that is conceptually simple, but leads to a more challenging joint inference problem.

Figure 7 illustrates the issue. For any given velocity parameters u and v , the model first predicts the mean $\mu(\phi)$ (from the sinusoidal function), and then predicts that the measured radial velocity $a(\phi)$ will come from a Normal distribution centered at $\mu(\phi)$. But whenever the predictive probability density falls outside the unambiguous interval, the model assigns probability to an event that cannot occur, because all measurements will always fall inside the unambiguous interval. In the example shown, the model places essentially all probability mass on an impossible event!

How can we fix the model? We still believe it is a good model for the *true* radial velocity—the only problem is aliasing. The obvious solution is to augment the model to also model the process of aliasing by mapping the probability density for any value a outside the unambiguous interval to its alias \bar{a} within the unambiguous interval. This leads to a model for the aliased radial velocity $\bar{a}(\phi)$ where only values inside the unambiguous interval have positive probability density, and each point \bar{a} accumulates the density of *each of its aliases* in the original model. Mathematically, this is described by the *wrapped normal* density function (Breitenberger 1963):

$$\mathcal{N}_w(\bar{a} \mid \mu, \sigma^2) = \sum_{k=-\infty}^{\infty} \mathcal{N}(\bar{a} + 2kV_{max} \mid \mu, \sigma^2)$$

$$\bar{a} \in (-V_{max}, V_{max}].$$

With this definition in place, our only change to the original model is to replace the normal likelihood in the uniform velocity model by the following wrapped normal likelihood:

$$\bar{a}(\phi) \sim \mathcal{N}_w(\mu(\phi), \sigma^2).$$

The resulting density is shown in the right panel of Figure 7: it looks like the original normal density but is positive only on the unambiguous interval and wraps around at $\pm V_{max}$.⁴

Inference

Our high-level approach is to use the wrapped normal density in place of the normal density in the uniform velocity model and then fit the parameters. However, the likelihood surface is considerably more complex than that of the simple linear model, so fitting the parameters is more difficult. This is again consistent with the general principle that joint inference is preferable, but it is usually more demanding computationally.

Problem: Multi-Modal Likelihood

Our goal is to find the parameters u and v that maximize the log-likelihood of the wrapped normal model. In the original model, the log-likelihood was extremely well behaved: it was convex (quadratic) and the global optimum could be found easily by a least squares fit. In contrast, Figure 8 shows examples of the (negative) log-likelihood surface of the wrapped normal model. We can see that it is non-convex and multi-modal. In cases with a lot of data (radial velocity measurements from many different pulse volumes at different angles) the surface typically has one high

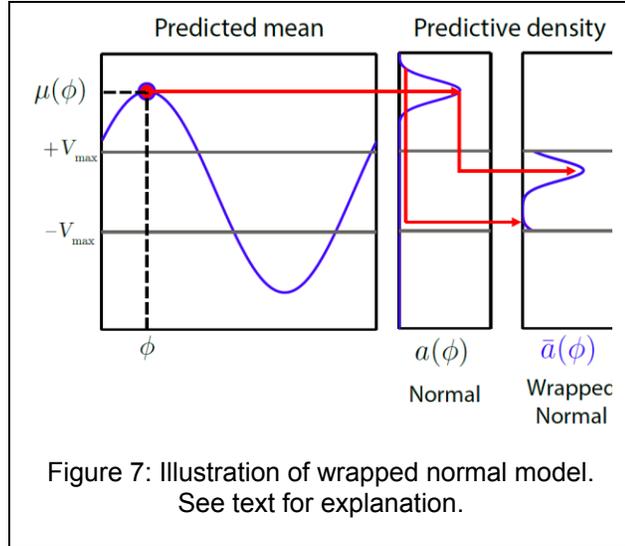
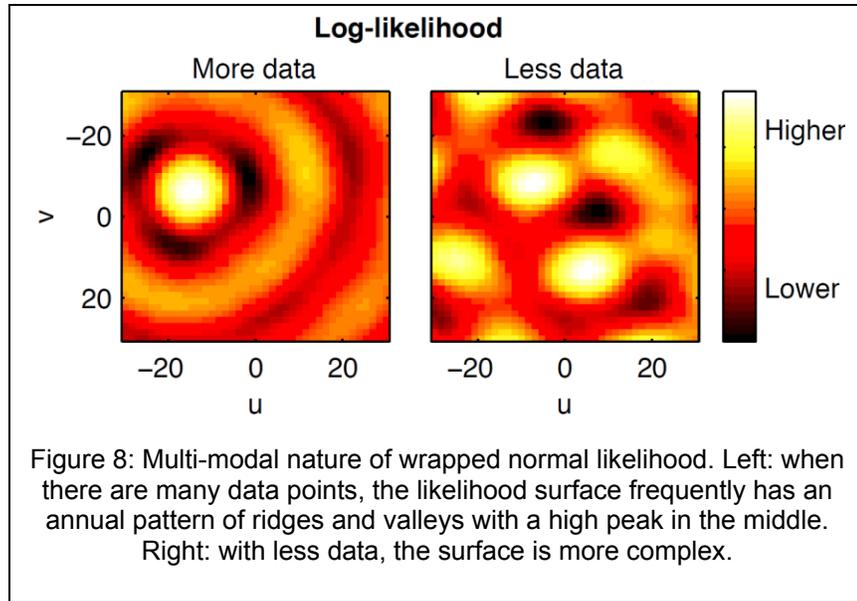


Figure 7: Illustration of wrapped normal model. See text for explanation.

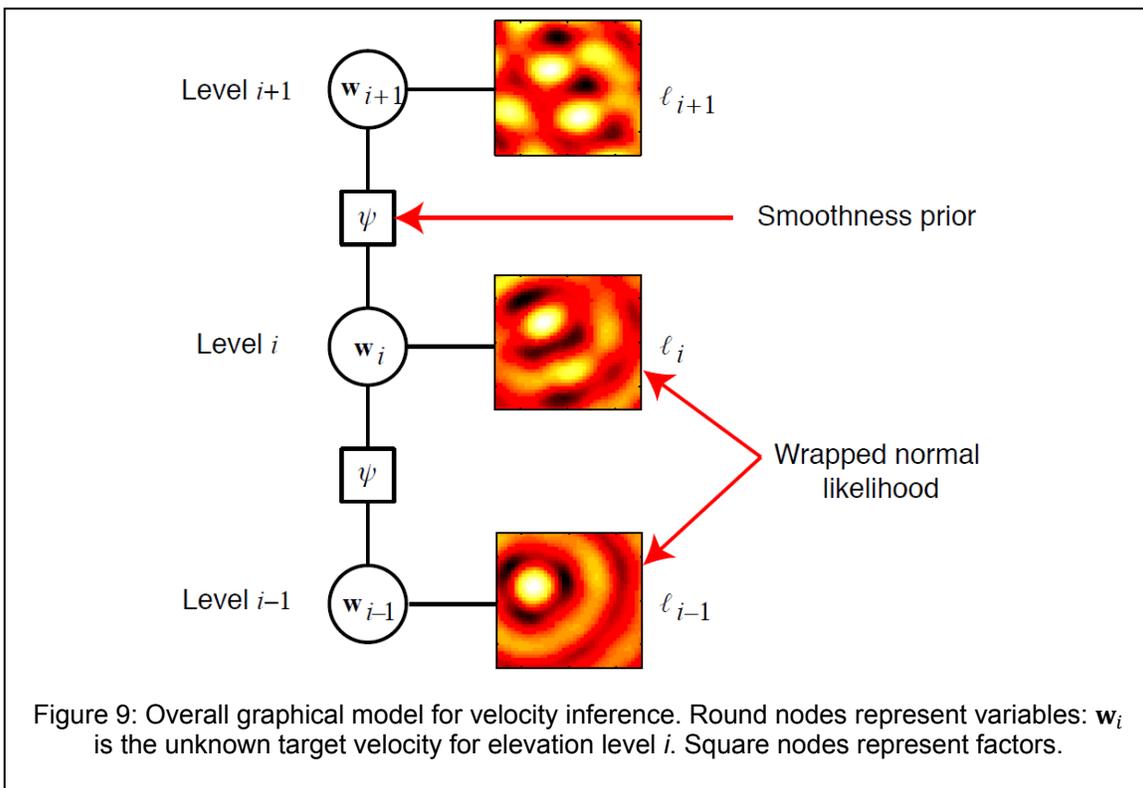
⁴When the variance is larger, the appearance of the wrapped normal density is less like its normal counterpart. As variance approaches infinity, the distribution approaches a uniform distribution on $(-V_{max}, V_{max}]$

peak centered near the “true” parameters surrounded by an annular pattern circular ridges and valleys. When there is less data, the pattern of the surface is more complex and has more local optima. Thus, any fitting method risks getting caught in a local optimum, which can cause the fitted parameters to be worse than those found by other methods despite having a better model.



A Graphical Model

Recall that our overall goal is to create elevation profiles by fitting the velocity parameters for each different elevation bin. To share information across elevations we use the wrapped normal likelihood within the graphical model whose factor graph is shown in Figure 9. The variables in the model are velocity vectors $\mathbf{w}_i = (u_i, v_i)$ for height levels $i \in \{1, 2, \dots\}$. In addition to the likelihood $\ell_i(\mathbf{w}_i)$ for each level, the model connects the parameters at adjacent elevations using a



weak Gaussian smoothness prior $\psi(\mathbf{w}_{i-1}, \mathbf{w}_i) = \exp(-\|\mathbf{w}_i - \mathbf{w}_{i-1}\|^2/2\tau^2)$. This has two benefits. The first is statistical: by sharing information, a level with strong evidence can guide the search to the correct mode in a level with weak evidence. For example, in Figure 9, the evidence gets progressively weaker from the bottom level to the top level. The correct mode for the top level is not visually apparent when looking at the local evidence alone, but the lower levels strongly support the choice of the mode that aligns well with peaks of those likelihood functions. The second benefit is computational: even when the evidence at a given layer is strong, the fitting method still has to find the correct mode, and evidence from adjacent layers can help initialize the search to do this quickly.

Approximate Inference by Expectation Propagation

Our goal is now to perform inference in the graphical model of Figure 9 to recover the posterior distribution of the \mathbf{w}_i variables. For example, the posterior mean of \mathbf{w}_i provides a point estimate of the velocity at level i .

Unfortunately, one cannot easily apply standard message-passing algorithms in this model to compute the exact posterior distribution. The wrapped normal likelihood functions $\ell_i(\mathbf{w}_i)$ are non-Gaussian continuous potentials. Because of this, there is no way to compactly represent the messages or efficiently perform the key subroutines of multiplication and marginalization of potentials. However, it is worth noting that if we had simply added the smoothness prior to the *original* uniform velocity model, all potentials would have remained Gaussian and exact inference would have remained efficient, because multiplication and marginalization can be done by simple parameter updates for Gaussian potentials (Murphy 2002). In fact, such a model is mathematically equivalent to the well-known Kalman filter.

However, because exact inference is intractable in our model, we instead adopt an approximate message passing scheme based on the Expectation Propagation (EP) algorithm (Minka 2001), which approximates each likelihood term by a Gaussian potential. An important aspect of EP is that the approximations are not made a priori, but *in the context of all other messages*—conceptually, EP focuses on finding a good approximation for the full posterior distribution instead of simply approximating the individual non-Gaussian potentials. Full details of the message passing scheme can be found in the conference paper (Sheldon et al. 2013).

Local Search

One additional detail of the message passing scheme is worth mentioning because it reveals that the basic computations performed by the overall approach have a simple and intuitive interpretation from the perspective of dealiasing.

As part of EP, we must solve the subproblem of finding a Gaussian approximation of a particular non-Gaussian distribution $f(\mathbf{w}_i)$ that is the product of the wrapped-normal likelihood $\ell_i(\mathbf{w}_i)$ with a prior over \mathbf{w}_i from adjacent levels. We do this using Laplace’s method—a variant of EP known as Laplace Propagation (Smola, Vishwanathan, and Eskin 2003)—for which we need to find the mode of $f(\mathbf{w}_i)$.

A very simple and intuitive local search procedure converges to an approximate mode of $f(\mathbf{w}_i)$ under a one-term approximation to the infinite sum in the wrapped normal density. The local search begins with an initial value for \mathbf{w}_i , and then alternates between two simple high-level steps. First, it uses the current value of \mathbf{w}_i to dealias the radial velocity measurements by selecting the alias closest to the predicted mean for each data point. Then, it refits \mathbf{w}_i by least squares using the current dealiased values. The search is very fast and usually converges in one or two

steps, and there is no measurable loss in overall performance compared with a much slower numerical approach using many more terms to approximate the infinite sum in the wrapped normal density. The overall algorithm then consists of very simple ingredients: local search as described here, together with standard operations on Gaussian potentials, such as those performed during Kalman filtering.

Other Approaches for Profiling with Aliased Data

While we are not aware of any previous work that jointly reasons about aliasing and velocity reconstruction, other authors have developed techniques to fit velocity profiles directly from aliased data through clever transformations of the data.

Tabary, Scialom, and Germann (2001) observed that the derivative of the function $\mu(\phi)$ from the uniform velocity model with respect to ϕ is also sinusoidal in ϕ and linear in u and v , so it is possible to fit the velocity parameters by first estimating the derivatives (for example, by fitting a locally linear model to the radial velocity data) and then solving a least squares problem. Gao et al. (2004) followed a similar approach using finite differences to estimate the derivatives. The advantage of both of these methods over the original model is that they are more robust to aliasing because the derivative (slope) of the sinusoid is unchanged if the curve is shifted up or down by aliasing. The *estimated* derivative may be contaminated if the curve crosses an aliasing boundary within the window used to make the estimate; however, such an error leads to very large values for the derivative that one can usually recognize and discard from the analysis.

While these methods are much more robust to aliasing than basic VAD, they are also subject to some limitations. In our conference paper (Sheldon et al. 2013), we extended and theoretically analyzed the gradient-based velocity azimuthal-display (GVAD) method of Gao and Droegemeier (2004). We showed that GVAD has the downside of increasing the variance of the parameter estimates by a factor related to the size of the window used to estimate the slope, and quantified precisely a tradeoff inherent in this approach: smaller windows lead to more noise in the fitting process, whereas larger windows lead to more aliasing errors. In the conditions frequently present during bird migration (low Nyquist velocity, increased fine-scale variability in the velocity field), it can be difficult to find a window size that is large enough to keep noise levels under control while still being reasonably robust to aliasing.

Experiments

We evaluated the accuracy of six different velocity profiling algorithms on volume scans collected by the KBGM radar station in Binghamton, NY during the month of September 2010. We screened an initial set of 351 scans (one per hour for each night of the month) to eliminate scans with clear non-biological targets (precipitation or ground clutter) within 37.5 km of the station. Of the remaining 142 scans, 102 operated in the clear-air mode with Nyquist velocity $V_{max} = 11.61 \text{ ms}^{-1}$ which is the lowest of all modes. We then used all pulse volumes with elevation angle $\rho \leq 5$ degrees and range $r \leq 37.5 \text{ km}$ to fit velocity vectors for each 20 m elevation bin below 6000 m. We use the following baseline algorithm:

- GVAD-0.01: The basic GVAD algorithm of Gao et al. (2004) with finite-difference derivative estimation.

We then tested the following five algorithms which incorporate different improvements to the baseline that culminate in our full EP message-passing algorithm:

- GVAD-0.10: An extension to GVAD to reduce variance by increasing the window size, at the cost of being slightly less robust to aliasing errors.
- GVAD+LOCAL: GVAD-0.10 followed by the local search procedure described above.
- GVAD+KALMAN: same as above, but followed by forward-backward message passing in a graphical model similar to the one in Figure 9 where the wrapped-normal likelihoods are approximated *a priori* by Gaussian potentials, so the model has the structure of a Kalman filter.
- EP: same as above, but followed by additional forward and backward passes using the EP message updates, with the approximate local search procedure used within EP.
- EP+HESS: same as above, but MATLAB's numerical optimizer and the more accurate five-term approximation to the infinite sum in the wrapped normal density are used in place of the local-search procedure.

Additional details about the experimental setup and the algorithms can be found in the conference paper (Sheldon et al. 2013).

Figure 10 shows the average RMSE and 95% confidence interval attained by each algorithm on the 142 test scans. All comparisons are highly significant ($p < 10^{-7}$; paired t-test) except GVAD+LOCAL vs. GVAD+KALMAN and EP vs. EP+HESS. It is clear from the results that each of the improvements we proposed makes a substantial improvement to performance. GVAD-0.10 is much better than GVAD-0.01, highlighting the importance of our extension to reduce variance by increasing the window size. By adding the simple dealias-and-refit local search procedure, GVAD+LOCAL performs much better than GVAD alone. Finally, by performing approximate Bayesian inference with the more accurate wrapped normal likelihood, EP performs significantly better than GVAD+LOCAL. The lesser performance of GVAD+KALMAN indicates that the graphical model structure alone does not explain the better performance of EP: the approach of approximating the wrapped normal likelihood in the context of the current posterior is important. Finally the performance of EP+HESS shows that EP loses no accuracy by using the approximate local search in the Laplace approximation.

By examining individual scans (not shown), we observed that the better models were generally in close agreement about the direction of travel, but the GVAD-based models seemed to underesti-

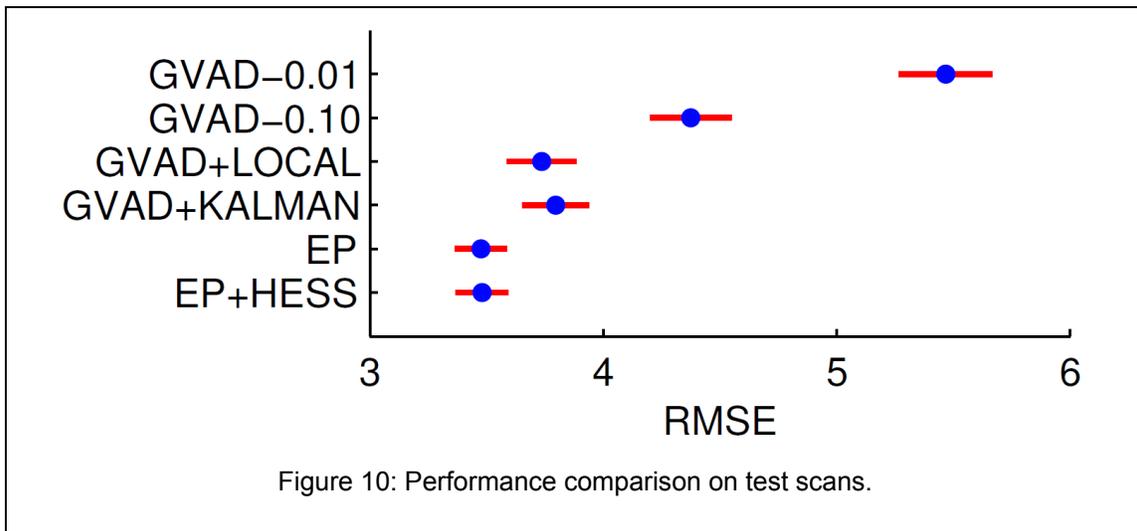
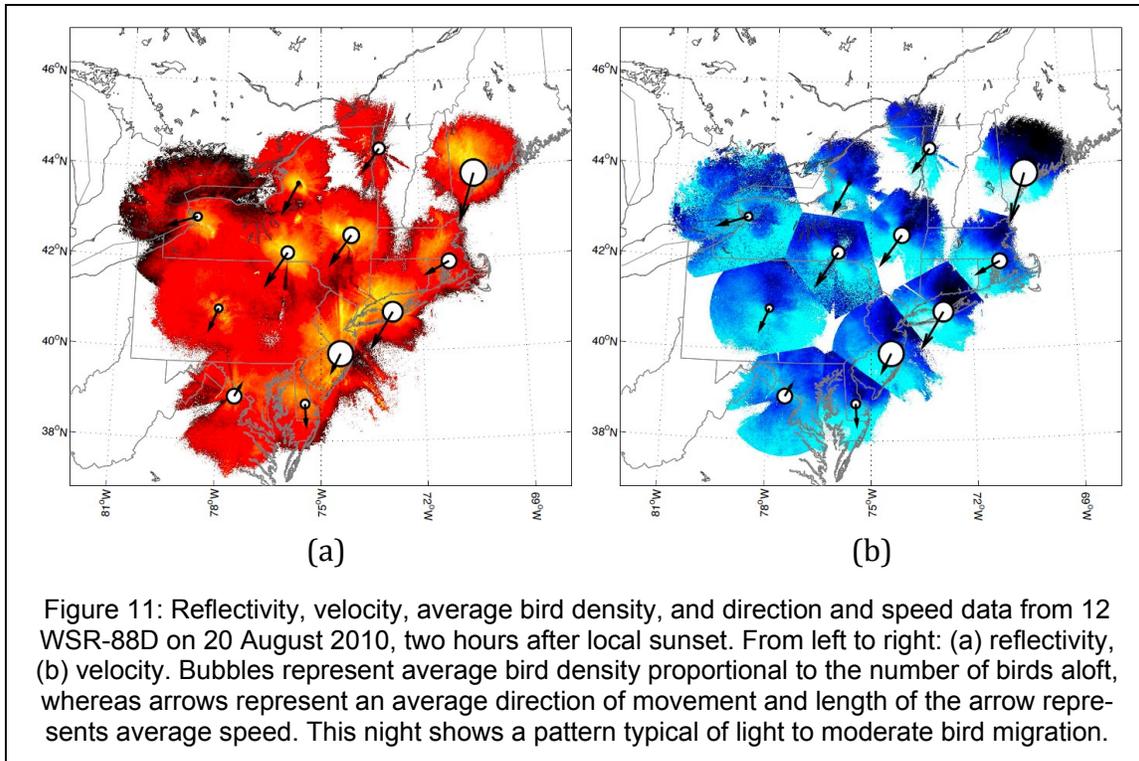


Figure 10: Performance comparison on test scans.



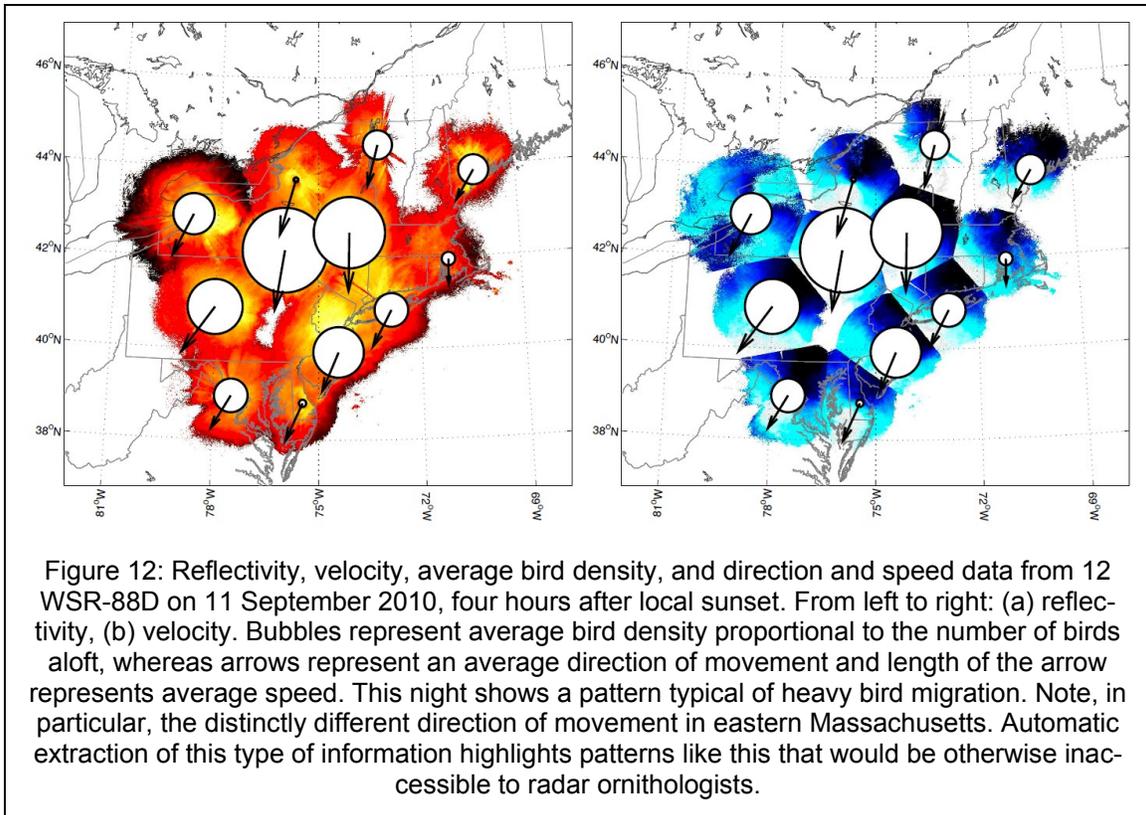
mate speed, which is likely due to not being completely robust to aliasing errors. The directions measured by the different algorithms were also in excellent agreement with human-labeled directions (mean error less than 10 degrees).

Examples of Bird Migration Across the Northeastern US

The application of this algorithm has yielded some exciting preliminary results and has the potential to uncover to new scientific knowledge about migration at large scales. We are currently applying this algorithm combined with other techniques to a database of 35,000 radar scans from 12 radar stations in the northeastern US spanning August to November in 2010 and 2011 (i.e., most of the fall migration seasons for those years). Our goal is to produce three key pieces of information for every hour of the night during this period for each radar station: a relative bird density and a direction and speed of movement for those birds.

By using our method to produce these three pieces of information, we can describe some of the fundamental features of nocturnal bird movements over this part of the US. This type of description of bird movements does not exist for the scale and with the level of hourly detail for any location in the US, so we are using our AI methods to reach new levels of visualization, interpretation, and analysis of bird migration. With this approach, we can produce graphics that show these basic attributes of nocturnal bird migration across the region, during the course of a night (e.g., to show when peak migration density occurs), and in many other ways.

Figures 11 and 12 show examples of output from our model for visualizing bird migration. The night of 20 August 2010 (Figure 11 (a,b)) was a good night for migration in many portions of the Coastal Plain from New Jersey north to southern Maine and inland through the foothills of the Appalachians and Catskills. A snapshot taken two hours after local sunset shows the magnitude, direction of movement, and speed for birds aloft and the raw reflectivity and velocity data from which that information was extracted. Most birds are moving SW or WSW at the locations where

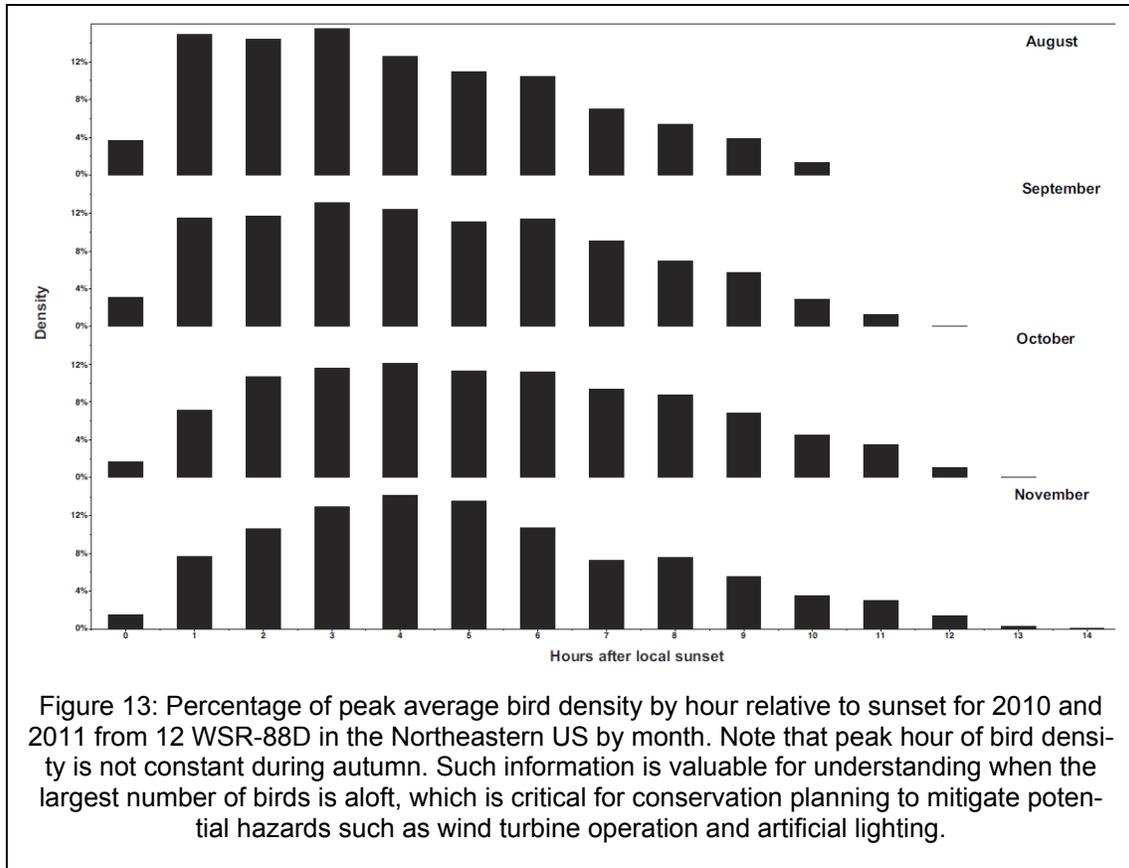


migration is occurring. Note, however, that there is variation among stations in the magnitude, direction, and speed of movements. This magnitude of movement is considered light to moderate migration. The night of 10 September 2010 (Figure 12 (a,b)) was an excellent night for migration across the same region. Heavy bird migration occurred, depicted here in a snapshot from four hours after sunset. Most movements were toward the SW or SSW, and the heaviest of those occurred in central New York.

Examining migration events in this way allows us to characterize bird movements, the conditions under which they occur, and the extent to which they vary in space and time. This information is invaluable for understanding not only how and where birds migrate, but also how protect valuable migration habitat, both in the air and on the ground. For example, visualizing the peak hour of bird density aloft over the 12 radar stations by month (Figure 13, a, b, c, d) shows us that densities tend to peak in the hours just after sunset, rather than closer to dawn. This suggests that in our study region, birds do not migrate for the whole night, but instead land after a few hours of flight. This type of characterization is much more powerful when done at large temporal and spatial scales, something made possible by the application of AI techniques and close collaborations between computer scientists and radar ornithologists.

Conclusion

Our approach to reconstructing the velocity fields of migrating birds detected by WSR-88D is a significant improvement over previous methods. Our algorithm will allow us to overcome a fundamental challenge of analyzing radar data to tap the potential information about bird migration available from the continent-scale WSR-88D network. By creating an AI tool to automate velocity data processing, we can extract information about bird migration more accurately and at a substantially larger scale than previously possible, and make advances in our knowledge of bird



movements for science and conservation. Of particular value is the potential for using this newly possible understanding of bird migration from our radar studies as a platform to monitor and assess ecosystem health. Many species of birds are excellent bioindicators (Holt and Miller 2011), and quantifying bird migration at regional and continental scales across 10-20 year intervals provides an invaluable opportunity to examine patterns of movements relative to changes in landscape and climate.

Collaborations between computer scientists and domain scientists in sustainability-related fields can yield powerful insights and advances in domains of mutual interest. In our case, uniting radar ornithology with machine learning research has yielded previously inaccessible and unavailable information about the behaviors of birds migrating over the northeastern United States. This powerful example can be a model for scaling up analyses to much larger spatial and temporal scales. Many different future analyses will benefit from having a robust quantification of bird densities and velocities aloft: for example, comparisons among radar stations in different parts of the US, comparisons with weather and climate data, and comparisons with species observations from eBird and other sources. Future work will expand on our existing dataset of 35,000 scans from 12 radar stations operating from 1 August to 30 November in 2010 and 2011 in the Northeast US to include the entire continent. Our approach also highlights additional algorithms necessary to continue the process of exploring this massive dataset. For example, a remaining challenge is to automate the screening of non-biological targets and to perform these analyses in real-time. A detailed understanding of the velocity field is a key first step toward this goal (Dokter et al. 2011); complete solution will likely involve applications of machine learning algorithms as well.

This article has described a case study in the AI techniques underlying automated analysis of bird migration at large scales using radar data. A fully automated continent-scale analysis of bird migration may be game changing in terms of understanding this global ecological process and protecting the system from growing anthropogenic threats. Recent advances in the new field of Computational Sustainability have the potential to make this a reality. We believe it is now possible to study migration at continent scale by integrating radar data together with complementary sources of information such as citizen science data from eBird and acoustic monitoring. Each of these data sources is available at the continent scale but provides only partial information about bird migration. The synthesis of these diverse data sources using novel analysis approaches will allow us to reconstruct, understand, and forecast nightly bird migrations at the continental scale.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant No. 1125228 and by the Leon Levy Foundation.

Literature Cited

Able, K.P. 1999. How Birds Migrate: Flight Behavior, Energetics, and Navigation. In *Gatherings of Angels: Migrating Birds and Their Ecology*, ed. K.P. Able, 11-27. Ithaca: Cornell University Press.

Alerstam, T. 1993. *Bird Migration*. Cambridge: Cambridge University Press.

Bonter, D.N., T. M. Donovan, and E. W. Brooks. 2009. Daily Mass Changes In Landbirds During Migration Stopover on the South Shore of Lake Ontario. *Auk* 124(1): 122-133.

Breitenberger, E. 1963. Analogues of the Normal Distribution on the Circle and the Sphere. *Biometrika* 50 (1): 81-88.

Buler, J. J., and R. H. Diehl. 2009. "Quantifying Bird Density During Migratory Stopover Using Weather Surveillance Radar." *IEEE Transactions on Geoscience and Remote Sensing* 47 (8): 2741-2751. doi:10.1109/TGRS.2009.2014463.

Buler, J. J., L. A. Randall, J. P. Fleskes, W. C. Barrow, T. Bogart, and D. Kluver. 2012. Mapping Wintering Waterfowl Distributions Using Weather Surveillance Radar. *PloS One* 7 (7): e41571. doi:10.1371/journal.pone.0041571.

Carey, J. 2012. Drones at Home. *Scientific American*, 307 (6), 44-45.

Crum, T. D., and R. L. Alberty. 1993. The WSR-88D and the WSR-88D Operational Support Facility." *Bulletin of the American Meteorological Society* 74 (9): 1669-1687.

Crawford, R. and T. Engstrom. 2001. Characteristics of Avian Mortality at a North Florida Television Tower: A 29-Year Study. *Journal of Field Ornithology* 72(3): p. 380-388.

Deppe, J.L. and J.T. Rotenberry. 2008. Scale-dependent habitat use by fall migratory birds: vegetation structure, floristics, and geography. *Ecological Monographs* 78(3): p. 461-487.

Dietterich, T. G., E. Dereszynski, R. Hutchinson, and D. Sheldon. 2012. Machine Learning for Computational Sustainability." In *Proceedings of the 2012 International Green Computing Conference (IGCC)*. IEEE Computer Society.

Dokter, A. M., F. Liechti, H. Stark, L. Delobbe, P. Tabary, and I. Holleman. 2011. Bird Migration Flight Altitudes Studied by a Network of Operational Weather Radars. *Journal of the Royal Society, Interface / the Royal Society* 8 (54): 30–43.

Doviak, R., and D. Zrnic. 1993. *Doppler Radar and Weather Observations*. Academic Press.

Eriksen, C.C., T.J. Osse, R.D. Light, T. Wen, T.W. Lehman, P.L. Sabin, J.W. Ballard, and A.M. Chiodi. 2001. Seaglider: A Long-Range Autonomous Underwater Vehicle for Oceanographic Research. *IEEE Journal of Oceanic Engineering* 26: 4.

Fink, D., W. M. Hochachka, B. Zuckerberg, D. W. Winkler, B. Shaby, M. A. Munson, G. Hooker, D. Sheldon, M. Riedewald, D. Sheldon, and S. Kelling. 2010. Spatiotemporal exploratory models for broad-scale survey data. *Ecological Applications* 20:2131-2147.

Fink, D., T. Damoulas, and J. Dave. 2013. Adaptive Spatio-Temporal Exploratory Models: Hemisphere-wide species distributions from massively crowdsourced eBird data. *AAAI Conference on Artificial Intelligence*, North America, jun. 2013.

Francis, C.M. 1995. How useful are recoveries of North American passerines for survival analyses? *Journal of Applied Statistics* 22(5): 1075.

Gao, J., K. K. Droegemeier, J. Gong, and Q. Xu. 2004. A Method for Retrieving Mean Horizontal Wind Profiles from Single-Doppler Radar Observations Contaminated by Aliasing. *Monthly Weather Review* (1982): 1399–1409.

Gao, J., and K. K. Droegemeier. 1990. A Variational Technique for Dealiasing Doppler Radial Velocity Data. *Journal of Applied Meteorology*: 934–940.

Gauthreaux, S. A., Jr., and C. G. Belser. 2003. Radar Ornithology and Biological Conservation. *Auk* 120 (2): 266–277.

Gauthreaux, S. A., Jr., C. G. Belser, and D. Van Blaricom. 2003. Using a Network of WSR-88D Weather Surveillance Radars to Define Patterns of Bird Migration at Large Spatial Scales. In *Avian Migration*, edited by P. Berthold, E. Gwinner, and E. Sonnenschein, 335–346. Springer-Verlag.

Holt, E. A. and S. W. Miller. 2011. Bioindicators: Using Organisms to Measure Environmental Impacts. *Nature Education Knowledge* 3(10): 8

Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and Environment* 6(5): 282–284.

Kelling, S., J. Gerbracht, D. F. Fink, C. Lagoze, W. Wong, J. Yu, T. Damoulas, and C. Gomes. 2012. A Human/Computer Learning Network to Improve Biodiversity Conservation and Research. *AI Magazine* 34 (1): 10.

Kunz, T. H., S. A. Gauthreaux Jr., N. I. Hristov, J. W. Horn, G. Jones, E. K. V. Kalko, R. P. Larkin, et al. 2008. Aeroecology: Probing and Modeling the Aerosphere. *Integrative and Comparative Biology* 48 (1): 1–11. doi:10.1093/icb/icn037.

Lack, D., and G. C. Varley. 1945. Detection of Birds by Radar. *Nature* 156: 446.

- Longcore, T., C. Rich, and S. A. Gauthreaux Jr. 2008. Height, guy wires, and steady-burning lights increase hazard of communication towers to nocturnal migrants: a review and meta-analysis. *Auk* 125(2): 486-493.
- Minka, T. 2001. A family of algorithms for approximate Bayesian inference. Ph.D. diss., Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Moore, F.R., R.J. Smith, and R. Sandberg. 2005. Stopover ecology of intercontinental migrants: En route problems and consequences for reproductive performance, in *Birds of Two Worlds – the ecology and evolution of migration.*, R. Greenberg and P. Marra, Editors, John Hopkins Press: Baltimore, MD. p. 251-261.
- Murphy, K. P. 2002. *Dynamic Bayesian Networks: Representation, Inference and Learning.* University of California, Berkeley.
- Newton, I. 2008. *The Migration Ecology of Birds.* Academic Press.
- Newton, I. 2008. Bird Migration, in *Handbook of the Birds of the World.* Vol. 13 Penduline-tits to Shrikes, J. del Hoyo, A. Elliott, and D.A. Christie, Editors. Lynx Edicions: Barcelona. p. 15-47.
- Packett, D.L. and J.B. Dunning. 2009. Stopover habitat selection by migrant landbirds in a fragmented forest-agricultural landscape. *Auk* 126(3): p. 579-589.
- Poon, H., and P. Domingos. 2007. Joint Inference in Information Extraction. In *AAAI*, 7:913–918.
- Schaub, M., F. Liechti, and L. Jenni. 2004. Departure of migrating European Robins, *Erithacus rubecula*, from a stopover site in relation to wind and rain. *Animal Behaviour* 67: p. 229-237.
- Sheldon, D., A. Farnsworth, J. Irvine, B. Van Doren, K. Webb, T. Dietterich, and S. Kelling. 2013. Approximate Bayesian Inference for Reconstructing Velocities of Migrating Birds from Weather Radar. *AAAI Conference on Artificial Intelligence*, North America, jun. 2013.
- Sherman, J., R. E. Davis, W. B. Owens, and J. Valdes, 2001. The autonomous underwater glider "Spray". *IEEE Journal of Oceanic Engineering* 26:437-446.
- Smola, A., S. V. N. Vishwanathan, and E. Eskin. 2003. Laplace Propagation. *Neural Information Processing Systems* (Section 5).
- Tabary, P., G. Scialom, and U. Germann. 2001. Real-Time Retrieval of the Wind from Aliased Velocities Measured by Doppler Radars. *Journal of Atmospheric and Oceanic Technology* 18 (6) (jun): 875–882. doi:10.1175/1520-0426(2001)018<0875:RTROTW>2.0.CO;2.
- Tøttrup, A.P., K. Thorup, and C. Rahbek., 2006. Patterns of change in timing of spring migration in North European songbird populations. *Journal of Avian Biology* 37: p. 84-92.
- Veltri, C. and D. Klem. 2005. Comparison of Fatal Bird Injuries from Collisions with Towers and Windows. *Journal of Field Ornithology* 76(2): p. 127-133.

Biographical Sketches



Andrew Farnsworth is a research associate in Information Science at the Cornell Lab of Ornithology. His primary research interest is in understanding the behavioral ecology of migration by using radar ornithology and acoustic monitoring to study nocturnal migrants. He began birding at age 5, and quickly developed an interest in bird migration.



Dan Sheldon is a professor in the Department of Computer Science at the University of Massachusetts, Amherst. His research focus is to develop algorithms to understand and make decisions about the environment using large data sets. He seeks to answer foundational questions (what are the general models and principles that underlie big data problems in ecology?) and also to build applications that transform large-scale data resources into scientific knowledge and policy.



Jeffrey Geevarghese is a master's student interested in machine learning and its applications. He is currently a research assistant with the Department of Computer Science, University of Massachusetts, Amherst. He has contributed to BirdCast by developing a dataset that summarizes bird migration information from weather radar data. He will extend this work to develop robust systems for automated analysis of weather radar data for studying bird migration patterns.



Jed Irvine is a Faculty Research Assistant at Oregon State University providing software development support to Tom Dietterich's Machine Learning group. His primary contribution to the BirdCast project has been the development of a web application to facilitate the rapid labeling of large quantities of radar images as being acceptable or not for the BirdCast data

pipeline. Jed's interest in bird migration stems from the many days spent hawk watching with his Dad when growing up on the east coast.



Benjamin Van Doren is a sophomore at Cornell University, studying Ecology and Evolutionary Biology. He is fascinated by bird migration and enjoys using computers to shed light on biological questions. His research into the phenomenon of morning flight was recognized last year when he won fifth place in the national Intel Science Talent Search in Washington, D.C.

Kevin Webb is a software engineer in the Information Science Program at the Cornell Lab of Ornithology. His technical areas of interest are geospatial programming, high performance and parallel computing, data modeling, and database centric application development. His primary objective is the application of software engineering to facilitate discovery in conservation and biological sciences.



Tom Dietterich is a professor in and the director of Intelligent Systems Research at Oregon State University. He leads the development and application of advanced machine learning algorithms to model bird migration. He is also President-Elect of the Association for the Advancement of Artificial Intelligence.



Steve Kelling is the director of Information Science at the Cornell Lab of Ornithology. He leads a team of ornithologists, computer scientists, statisticians, application developers, and data managers to develop programs, tools, and analyses to gather, understand, and disseminate information on birds and the environments they inhabit.