# A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model

J. C. Lin,[1] C. Gerbig,[1] S. C. Wofsy,[1] A. E. Andrews,[2] B. C. Daube,[1] K. J. Davis,[3] and C. A. Grainger[4]

Received 8 November 2002; revised 3 April 2003; accepted 23 April 2003; published 19 August 2003.

[1] We introduce a tool to determine surface fluxes from atmospheric concentration data in the midst of distributed sources or sinks over land, the Stochastic Time-Inverted Lagrangian Transport (STILT) model, and illustrate the use of the tool with $CO_2$ data over North America. Anthropogenic and biogenic emissions of trace gases at the surface cause large variations of atmospheric concentrations in the planetary boundary layer (PBL) from the "near field," where upstream sources and sinks have strong influence on observations. Transport in the near field often takes place on scales not resolved by typical grid sizes in transport models. STILT provides the capability to represent near-field influences, transforming this noise to signal useful in diagnosing surface emissions. The model simulates transport by following the time evolution of a particle ensemble, interpolating meteorological fields to the subgrid scale location of each particle. Turbulent motions are represented by a Markov chain process. Significant computational savings are realized because the influence of upstream emissions at different times is modeled using a single particle simulation backward in time, starting at the receptor and sampling only the portion of the domain that influences the observations. We assess in detail the physical and numerical requirements of STILT and other particle models necessary to avoid inconsistencies and to preserve time symmetry (reversibility). We show that source regions derived from backward and forward time simulations in STILT are similar, and we show that deviations may be attributed to violation of mass conservation in currently available analyzed meterological fields. Using concepts from information theory, we show that the particle approach can provide significant gains in information compared to conventional gridcell models, principally during the first hours of transport backward in time, when PBL observations are strongly affected by surface sources and sinks. *INDEX TERMS:* 0315 Atmospheric Composition and Structure: Biosphere/atmosphere interactions; 0368 Atmospheric Composition and Structure: Troposphere—constituent transport and chemistry; 3307 Meteorology and Atmospheric Dynamics: Boundary layer processes; 3322 Meteorology and Atmospheric Dynamics: Land/atmosphere interactions; *KEYWORDS:* Lagrangian particle dispersion models, boundary layer transport, land-atmosphere exchange, carbon fluxes, trajectory modeling, global change science

**Citation:** Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger, A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.*, *108*(D16), 4493, doi:10.1029/2002JD003161, 2003.

## 1. Introduction

[2] Quantitative understanding of the sources and sinks of chemically and radiatively active trace gases is essential in order to assess the human impact on the environment [*Heimann and Kasibhatla*, 1999; *Mooney et al.*, 1987]. Observations of atmospheric concentrations provide the basic data for analyzing sources and sinks at the surface. However, the influences of surface fluxes are strongly modified by atmospheric transport, which tends to integrate over regional and continental scales, the relevant spatial scales needed to formulate policy responses to many of the anthropogenic perturbations. Thus accurate modeling of atmospheric transport is critical to quantitatively link observed tracer distributions to surface emissions and address issues of environmental concern.

[3] Trace gas concentrations in the Planetary Boundary Layer (PBL) are especially sensitive to surface fluxes, but

[1]Department of Earth and Planetary Sciences and Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA.
[2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[3]Department of Meteorology, Pennsylvania State University, University Park, Pennsylvania, USA.
[4]Department of Atmospheric Sciences, University of North Dakota, Grand Forks, North Dakota, USA.


0148-0227/03/2002JD003161$09.00

modeling atmospheric transport to interpret observations in the PBL is inhibited by two main challenges: (1) "near-field" variability in concentration data associated with inhomogeneities of the distribution of surface fluxes and (2) inadequate representation of PBL dynamics and transport in most models. The term near field is used loosely here to refer to the surface with which PBL air has come in contact before arrival at the observation location. The globally averaged ventilation time for the PBL is ~4 days [*Cotton et al.*, 1995]; hence the near-field domain affected by PBL processes may extend over $10^2$–$10^3$ km, i.e., regional or even continental scale. There is a strong need for diagnostic models that can be used to determine surface fluxes from observations in the PBL.

[4] Lagrangian particle dispersion models have the potential to address the above challenges by simulating turbulence and capturing subgrid scale transport. Particle models simulate both advection and dispersion in the PBL by explicitly incorporating turbulent velocity statistics in the trajectories of tracer particles [*Stohl*, 1998; *Zannetti*, 1990], expanding beyond traditional mean wind trajectory models that assume air parcels to be conserved entities [*Stohl and Wotawa*, 1993]. Particle locations are distributed without gridcell restrictions and can thus capture fine structures due to small-scale inhomogeneities in source distributions that are unresolved by gridded transport models.

[5] We introduce here a receptor-oriented framework for inferring surface sources or sinks of trace gases from atmospheric data in the PBL, the Stochastic Time-Inverted Lagrangian Transport (STILT) model. STILT simulates upstream influences on the observation location (receptor) by following the evolution of a particle ensemble backward in time. Analyzed wind fields are interpolated to the location of each particle, and turbulent motions are simulated by a Markov chain process based statistically on observed meteorological parameters.

[6] We examine in detail the physical and numerical requirements for STILT applicable to any particle model intended for use in interpreting atmospheric tracer observations. Empirical assessment of time reversibility provides key insight into physical and numerical inconsistencies in model frameworks and associated input data. We also present a quantitative estimate of the gain in information obtained from a particle approach over a standard Eulerian gridcell approach by employing concepts from information theory. The value of STILT as a near-field tool is illustrated by interpreting atmospheric data from the $CO_2$ Budget and Rectification Airborne (COBRA) study over North America.

## 2. Theoretical Background: Lagrangian Particle Dispersion Models

[7] Lagrangian particle transport models are based upon the assumption that turbulent diffusion can be modeled as a Markov chain, first suggested by *Obukhov* [1959] and *Smith* [1968]. *Hanna* [1979] has shown from Eulerian and Lagrangian measurements that the Markov assumption is reasonable, i.e., the particle velocity vector $\mathbf{u}$ can be decomposed into a mean component $\overline{\boldsymbol{u}}$ and a turbulent component $\mathbf{u}'$, with the turbulent component following the relation:

$$\mathbf{u}'(t+\Delta) = R(\Delta t)\mathbf{u}'(t) + \mathbf{u}''(t), \tag{1}$$

where $\mathbf{u}''$ is a random vector, $R$ is an autocorrelation coefficient, and $\Delta t$ is the time step. The autocorrelation coefficient $R$ in equation (1) follows an exponential form $R(\Delta t) = \exp(-\Delta t/T_{\mathrm{Li}})$, where $T_{\mathrm{Li}}$ is the Lagrangian timescale in the $i$ direction ($i$=$u$, horizontal; $i$=$w$, vertical). $T_{\mathrm{Li}}$ is a decorrelation timescale that determines the degree to which the particle's movement behaves like a random walk ($T_{\mathrm{Li}} = 0$) or like advection by mean wind (large $T_{\mathrm{Li}}$). The random velocity $\mathbf{u}''$ in equation (1) is defined by the relation:

$$\mathbf{u}'' = \boldsymbol{\lambda}\left[1 - R^2(\Delta t)\right]^{1/2}, \tag{2}$$

where $\boldsymbol{\lambda}$ is usually drawn from a Gaussian distribution with a mean of 0 and a standard deviation, $\sigma_i$, which characterizes the spread in the random velocity. The properties emerging from an ensemble of particle trajectories obtained from integrating equation (1) are used to characterize the turbulent flow.

[8] Particles transported forward in time provide a straightforward way for quantifying the effect an emission source has on downstream concentrations [*Lamb et al.*, 1979; *Ryall et al.*, 2001]. The total amount of tracer emitted is divided equally into the number of particles starting from the emission location, and the particle density at a specified location (receptor) directly yields the tracer concentration. Provided the particle transport is time reversible, the backward time run starting from the receptor results in the same particle number at the emission source. Hence emissions from upstream regions with more particles cause bigger changes in concentrations at the receptor. The particle density from backward time simulations provides the "influence density": the change in tracer concentration at the receptor in response to fluxes at the locations and times where the particles are found in the time-reversed model [*Uliasz and Pielke*, 1990]. Backward time particle simulations thus enable implementation of a "receptor-oriented framework" [*Uliasz and Pielke*, 1990] that defines upstream influences on tracer observations at the receptor. The backward time particle locations map out the influence function $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$, which quantitatively links sources and sinks $S(\mathbf{x}, t)$ to concentrations $C(\mathbf{x}_r, t_r)$ of a conserved tracer at a receptor located at $\mathbf{x}_r$ and time $t_r$, following *Holzer and Hall* [2000] and *Zannetti* [1990]:

$$C(\mathbf{x}_r,t_r) = \int_{t_0}^{t_r} dt \int_V d^3x I(\mathbf{x}_r,t_r \mid \mathbf{x},t) S(\mathbf{x},t) + \int_V d^3x I(\mathbf{x}_r,t_r \mid \mathbf{x},t_0) C(\mathbf{x},t_0). \tag{3}$$

$I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ has units of density (1 volume$^{-1}$), and measurements of $C(\mathbf{x}_r, t_r)$ are given in mixing ratio units (e.g., ppm), and the source/sink $S(\mathbf{x}, t)$ has units of mixing ratio per unit time (e.g., ppm h$^{-1}$). $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ specifies the fraction of a fluid element found at $\mathbf{x}_r$ and time $t_r$, given that the fluid element was at $\mathbf{x}$ and time $t$; it is analogous to the Green's function in the solution of the tracer continuity equation [*Holzer and Hall*, 2000]. The first term on the RHS of equation (3) denotes the concentration change at the receptor from sources/sinks in domain $V$ during the time

interval between initialization time $t_0$ and $t_r$. The second term refers to the contribution to the receptor concentration from advection of tracers from the initial tracer field $C(\mathbf{x}, t_0)$.

[9] The connection between influence function $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ and particle locations is formally established as follows. $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ is represented by the density $\rho(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ of particles at $(\mathbf{x},t)$ which were transported backward in time from $(\mathbf{x}_r,t_r)$, normalized by the total particle number $N_{\text{tot}}$:

$$I(\mathbf{x}_r, t_r \mid \mathbf{x}, t) = \frac{\rho(\mathbf{x}_r, t_r \mid \mathbf{x}, t)}{N_{\text{tot}}} = \frac{1}{N_{\text{tot}}} \sum_{p=1}^{N_{\text{tot}}} \delta\big(\mathbf{x}_p(t) - \mathbf{x}\big), \quad (4)$$

where the delta function in equation (4) simply represents the presence or absence of particle $p$ at location $\mathbf{x}$. The fields of $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ and $S(\mathbf{x}, t)$, continuous in space and time, are in practice resolved only at finite resolution with a discrete volume ($\Delta x$, $\Delta y$, $\Delta z$) and finite time interval ($\tau$). The influence function, integrated over the finite volume and time elements and applying equation (4), is given by:

$$\begin{aligned}
&\int_{t_m}^{t_m+\tau} dt \int_{x_i}^{x_i+\Delta x} dx \int_{y_j}^{y_j+\Delta y} dy \int_{z_k}^{z_k+\Delta z} dz\, I(\mathbf{x}_r, t_r \mid \mathbf{x}, t) \\
&= \frac{1}{N_{\text{tot}}} \int_{t_m}^{t_m+\tau} dt \int_{x_i}^{x_i+\Delta x} dx \int_{y_j}^{y_j+\Delta y} dy \int_{z_k}^{z_k+\Delta z} dz\, \rho(\mathbf{x}_r, t_r \mid \mathbf{x}, t) \\
&= \frac{1}{N_{\text{tot}}} \sum_{p=1}^{N_{\text{tot}}} \Delta t_{p,m,i,j,k}. \quad (5)
\end{aligned}$$

Thus the time- and volume-integrated influence function is simply quantified by tallying $\Delta t_{p,m,i,j,k}$, the total amount of time each particle $p$ spends in a volume element $i,j,k$ over time step $m$. Equation (5) provides source-receptor matrix elements that link sources at finite temporal and spatial resolutions directly to receptor concentrations through equation (3).

[10] The connection between "surface" fluxes $F(x,y,t)$ in units of (moles time$^{-1}$) to $S(\mathbf{x},t)$, a volume source/sink in units of (mixing ratio time$^{-1}$) can be established by diluting the surface tracer flux into an atmospheric column of height $h$, below which turbulent mixing is strong enough to thoroughly mix the surface signal during a time step:

$$S(\mathbf{x}, t) = \begin{cases} \dfrac{F(x, y, t) m_{\text{air}}}{h \bar{\rho}(x, y, t)} & \text{for} \quad z \le h \\ 0 & \text{for} \quad z > h \end{cases}, \quad (6)$$

where $\bar{\rho}$ is the average density below $h$ and $m_{\text{air}}$ is the molar mass of air. Given discretized surface fluxes $F(x_i, y_j, t_m)$, "footprint" elements $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$ linking the surface fluxes to concentration changes at the receptor $\Delta C_{m,i,j}(\mathbf{x}_r, t_r)$ can be derived as follows. Equation (6) is substituted for $S(\mathbf{x},t)$ in the first term on the RHS of equation (3) and integrated over discrete volume and time elements yielding

$$\begin{aligned}
&\Delta C_{m,i,j}(\mathbf{x}_r, t_r) \\
&= \left[ \frac{m_{\text{air}}}{h \bar{\rho}\big(x_i, y_j, t_m\big)} \int_{t_m}^{t_m+\tau} dt \int_{x_i}^{x_i+\Delta x} dx \int_{y_j}^{y_j+\Delta y} dy \int_{0}^{h} dz\, I(\mathbf{x}_r, t_r \mid \mathbf{x}, t) \right] \\
&\cdot F\big(x_i, y_j, t_m\big) = f\big(\mathbf{x}_r, t_r \mid x_i, y_j, t_m\big) F\big(x_i, y_j, t_m\big). \quad (7)
\end{aligned}$$

For surface flux $F(x_i, y_j, t_m)$ in units of ($\mu$mol m$^{-2}$ s$^{-1}$) and $\Delta C$ in (ppm), $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$ is in units of (ppm ($\mu$mol$^{-1}$ m$^2$ s)): given a unit surface flux of 1 $\mu$mol m$^{-2}$ s$^{-1}$ at $(x_i, y_j, t_m)$ persisting over a time interval $\tau$, $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$ gives the concentration change $\Delta C$ in ppm at the receptor. The $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$ can be derived from particle locations by combining equation (5) and the expression in brackets in equation (7):

$$f\big(\mathbf{x}_r, t_r \mid x_i, y_j, t_m\big) = \frac{m_{\text{air}}}{h \bar{\rho}\big(x_i, y_j, t_m\big)} \frac{1}{N_{\text{tot}}} \sum_{p=1}^{N_{\text{tot}}} \Delta t_{p,i,j,k}. \quad (8)$$

[11] Backward time particle simulations enable adoption of a receptor-oriented framework like equation (3) in an efficient way. A single backward time particle ensemble starting at a given $(\mathbf{x}_r, t_r)$ provides both the spatial and temporal dependence of $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$. The alternative approach of modeling transport forward in time from potential upstream influence areas incurs a much larger computational cost; no a priori method exists to decide which sources would influence the receptor, and forward time transport starting from the entire model domain $V$ needs to be simulated (Figure 1) for each time step between $t_0$ and $t_r$. To resolve a 1000 km $\times$ 1000 km domain at 50 km resolution for 2 days at 30 min resolution, the forward time simulation incurs a computation effort that is $10^4$ times larger than the backward time simulation.

[12] Forward and backward time simulations can be interpreted as selecting for different subsets of the "trajectory space": the collection of trajectories from all starting locations at all time steps within a temporal window. A forward simulation starting at source region $(x_i',y_i',z_i')$ selects for the subset of particle trajectories with $(x_i',y_i',z_i')$ as the starting point, while the backward simulation starting at receptor location $\mathbf{x}_r = (x_r,y_r,z_r)$ selects for the subset of particle trajectories arriving (viewed forward in time) at the receptor. The upstream influence on air at the receptor is determined precisely by air parcels that arrive at the receptor, so the backward time run selects for the appropriate air parcels in a single simulation. In contrast, only a small fraction of forward time particles starting at $(x_i',y_i',z_i')$ arrives at the receptor, so multiple simulations starting at different $(x_i',y_i',z_i')$ are needed to identify all trajectories that reach the receptor. Demonstration of time reversibility in the model would then show that a single backward time simulation generates the same $I(\mathbf{x}_r,t_r|\mathbf{x},t)$ as numerous forward time runs, leading to significant savings in computational effort.

[13] Lagrangian particle models have been used by numerous researchers in the forward time sense to model atmospheric transport [*Fay et al.*, 1995; *Luhar and Britter*, 1989; *Ryall et al.*, 2001; *Stohl et al.*, 1998]. Few studies to date have used backward time particle dispersion models to interpret and understand atmospheric tracer observations. *Uliasz and Pielke* [1990] pioneered the use of backward time particle models to derive influence areas that provide quantitative measures of the effects of different emissions on the air quality at receptor sites. *Uliasz* [1996] compared methylchloroform measured at receptor sites with values from the dispersion model, we are not aware of other studies

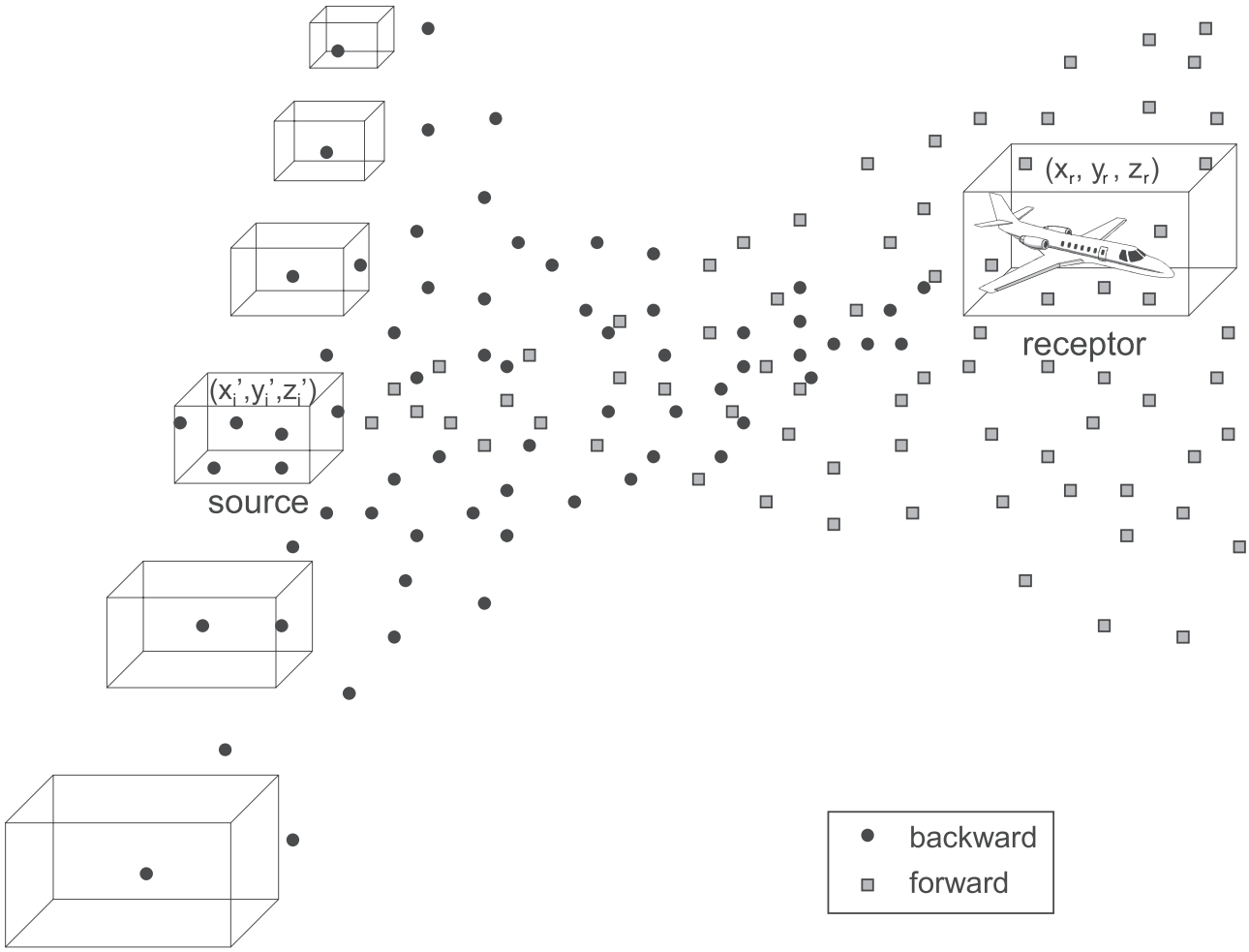


**Figure 1.** Comparison of backward time and forward time simulations. A single backward time release of particles marks out the potential source region that influences the receptor, yielding the spatial and temporal dependence of the influence function $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$, while numerous forward time runs from the entire model domain at multiple time steps are necessary. An empirical test for assessing time reversibility of the STILT model was carried out by simulating particles back in time from a box centered at the receptor location $\mathbf{x}_r = (x_r,y_r,z_r)$. Then numerous forward time runs were conducted, starting from each potential source region $(x_i',y_i',z_i')$. The number of particles from the backward run which end up in the source box is compared with the number from the forward run that is found in the receptor box, after taking into account the differences in air density at the source and the receptor. If the backward run yields similar information as the forward run, then the particle numbers should be comparable.

coupling observations with a backward time turbulent particle transport to derive source information.

## 3. Physical Requirements for Realistic Simulation of Source-Receptor Relationships

[14] Realistic simulations of transport and upstream influence by particle models must satisfy the following criteria: (1) well mixedness; (2) simulation of the close interaction between windshear and vertical turbulence; (3) high temporal resolution to resolve the decay in the autocorrelation of $\mathbf{u}'$; (4) consistent representation of particles as air parcels of equal mass in both the mean and turbulent transport components of the model.

### 3.1. Well Mixedness

[15] The well-mixed criterion specifies that models of turbulent dispersion must retain an initially well-mixed particle distribution and not spuriously unmix particles [*Thomson*, 1987], a consequence of the Second Law of Thermodynamics. The well-mixed criterion can be violated due to inconsistencies in both the turbulent and mean wind transport components. Turbulent parameterizations that inadequately account for the "drift velocity" associated with spatial gradients in turbulence variables cause particles to become trapped in low-turbulence regions [*Legg and Raupach*, 1982]. Mean wind fields that violate mass conservation also bring about the unmixing of particles. Mass conservation requires the amount of mass entering a gridbox to equal the one exiting it. Wind fields that spuriously create or destroy atmospheric mass lead to artificial dilution or concentration of particles and cause deviation from well mixedness (Figure 2).

[16] *Thomson* [1987] has theoretically demonstrated the equivalence of the well-mixed criterion to time reversibility. Thus the influence function $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ derived from an

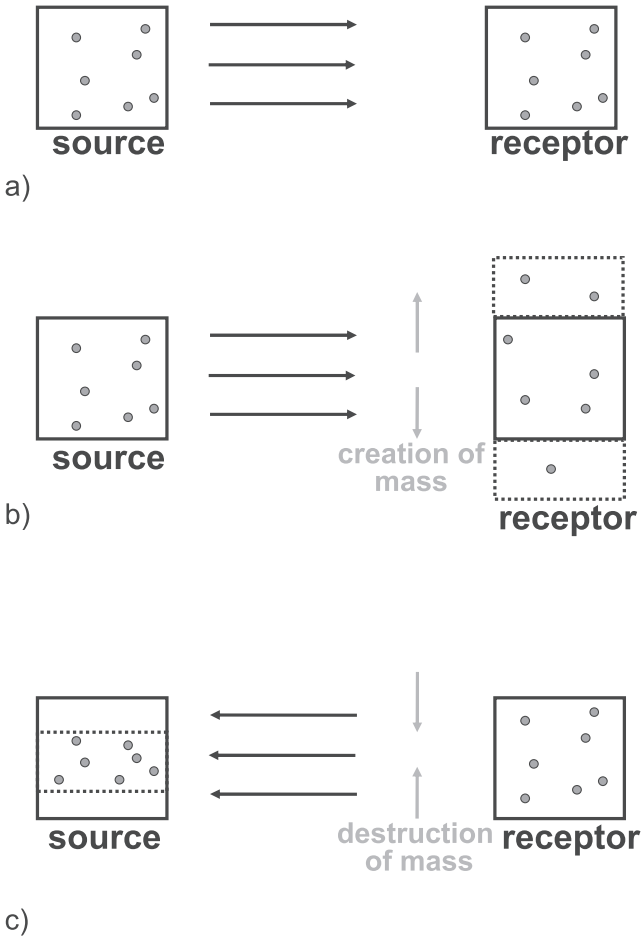


**Figure 2.** Schematic illustrating the unmixing of particles and time asymmetry introduced by mass violation in the driving wind fields. (a) Mass-conserving, one-dimensional flow. (b) Creation and (c) destruction of mass cause dilution and concentration of particles, respectively, and the effect on particles depends on the direction of time, leading to asymmetry between forward and backward time simulations.

ensemble of particle trajectories running forward in time is expected to be the same as $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ derived backward in time only if the well-mixed criterion is satisfied. Problems causing deviation from well mixedness introduce asymmetries into the forward and backward time simulations. Inconsistencies in turbulence parameterization which lead to particles being trapped in low-turbulence regions cause particle transport to be time irreversible: if forward particles were started in a low-turbulence region at the source, and the receptor was in a high-turbulence region, then few forward time particles would reach the receptor, whereas lots of backward time particles would become trapped in low-turbulence regions and end up at the source. The time asymmetry arising from mass violation can be traced to the fact that the effect of mass violation on particles, whether dilution or concentration, depends on the direction of time. The time asymmetry can be understood by considering a simple one-dimensional flow, shown in Figure 2. Forward in time, mass creation causes unphysical dilution of particles and fewer reach the receptor box as a result. Under time reversal, mass destruction causes the particles traveling backward from the receptor to be concentrated, and a greater number of particles reach the source box.

[17] Hence violation of the well-mixed criterion causes both forward and backward time results to be wrong. Comparison of forward and backward results is thus an effective way to reveal deficiencies in the model and has been conducted for the STILT model (see below).

### 3.2. Simulation of Interactions Between Windshear and Vertical Turbulence

[18] Dispersion is the result of the close coupling between the mean wind field and turbulent deviations from the mean field. The rapid vertical transport due to turbulent vertical velocity $w'$ displaces air parcels to different altitudes in the PBL, and the shear in mean wind causes air parcel trajectories at different altitudes to diverge from one another. Lagrangian particle models usually split the transport operator acting on particles into the mean and turbulent components. Implementation of transport operator splitting without capturing this tight coupling results in unrealistic simulations; particles are transported up and down in the turbulence module without being advected horizontally by mean wind, while particles are advected by the mean transport module without being redistributed vertically.

### 3.3. High Temporal Resolution to Resolve the Decay of the Autocorrelation in $\mathbf{u}'$

[19] Proper simulation of the Markov process given by equation (1) depends upon adequate resolution of the decay in autocorrelation of $\mathbf{u}'$ from the previous time step. The variable which dictates the timescale over which the autocorrelation persists is $T_L$, the Lagrangian timescale. A time step which is too large ($\geq T_L$) means that the particle motion is not affected by enough random velocity components, as the same $\mathbf{u}'$ persists through the entire time step instead of decaying. *Wilson and Zhuang* [1989] have suggested that a time step that is smaller than 10% of the local $T_L$ should be used.

### 3.4. Consistent Representation of Particles as Air Parcels of Equal Mass

[20] Particles advected by mean wind vectors are treated as air parcels of equal mass; they are concentrated and diluted while being transported lower and higher in the atmosphere, in accordance with the large vertical gradients in air density. The model's turbulent parameterization must take into account the atmosphere's vertical density gradient to maintain consistent treatment of the particles. Neglecting the density gradient will systematically underestimate tracer concentrations near the surface and overestimate concentrations at higher altitudes [*Stohl and Thomson*, 1999].

## 4. Model Description: STILT Model

[21] The STILT model is built upon source code from the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) system [*Draxler and Hess*, 1998], using the mean advection scheme from HYSPLIT but employing a different turbulent module. The HYSPLIT mean trajectories have been evaluated by comparison with neutral balloon paths [*Draxler and Hess*, 1998]. The turbulence parameterization is described below.

[22] The horizontal Lagrangian timescale $T_{\mathrm{Lu}}$ is assumed to be 3 hours velocity [*Draxler and Hess*, 1998]. The

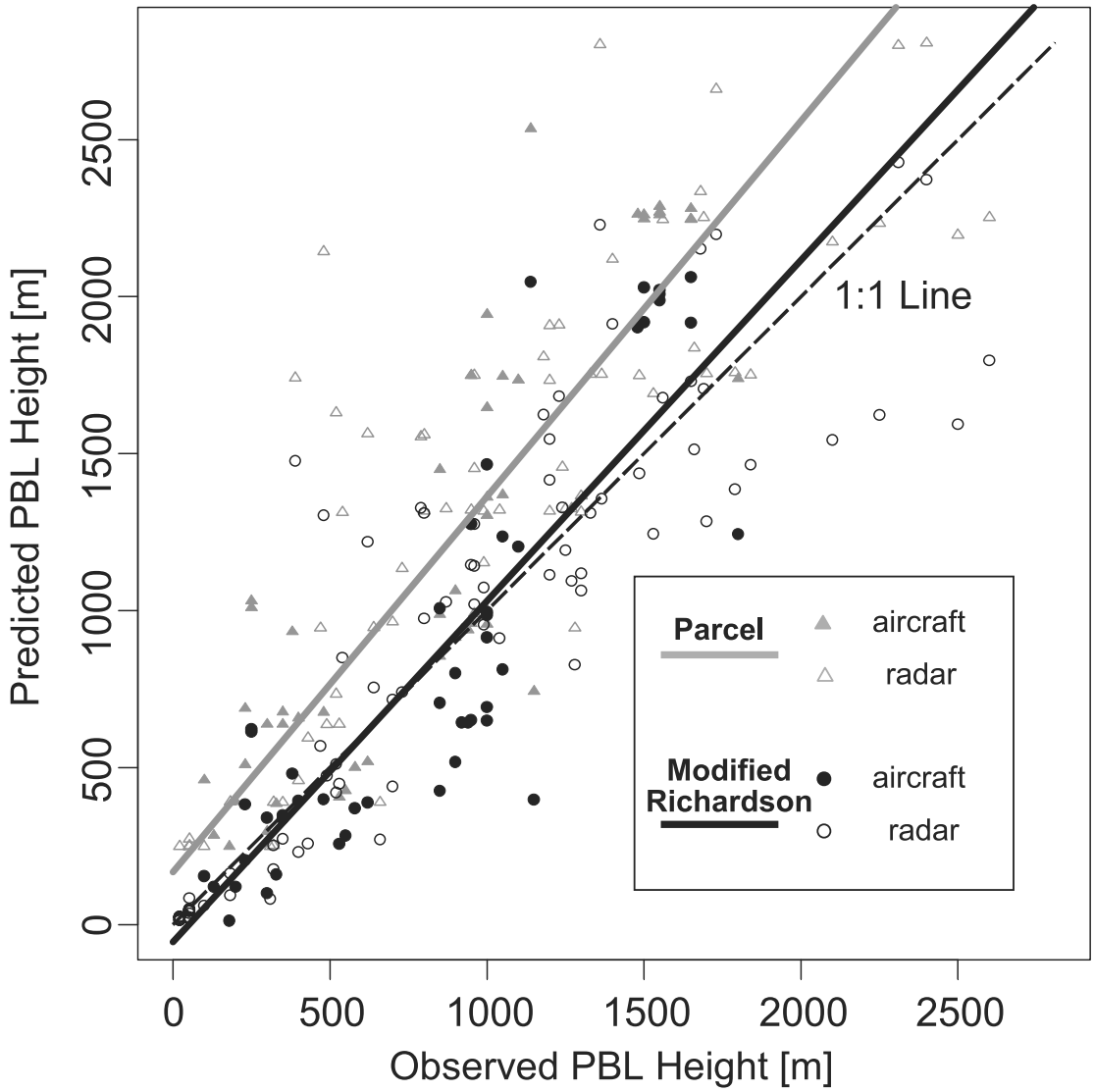


**Figure 3.** Comparison between observed and predicted values of PBL height using two different parameterizations: (1) the original parcel method used in HYSPLIT; result of orthogonal distance regression: $PBL_{pred} = 1.20PBL_{obs} + 167.5$; and (2) the new modified Richardson number method that yields predictions closer to observed values; result of orthogonal distance regression: $PBL_{pred} = 1.09PBL_{obs} - 55.1$. Solid symbols refer to observations derived from aircraft tracers, while open symbols denote observations using radar in Wisconsin.

parameters $T_{Lw}$ and $\sigma_w$ controlling vertical turbulence are represented, following *Hanna* [1982], as functions of the PBL height, roughness length, Monin-Obukhov length, convective velocity scale, and frictional velocity. The parameterization for PBL height, necessary because the general unavailability of PBL height in meteorological fields, has been modified from the simple "parcel method" in HYSPLIT, which specifies the PBL height as the level where the virtual potential temperature is greater than the surface value by 2°K. The new parameterization is a modified Richardson number method that generalizes to unstable, neutral, and stable conditions [*Vogelezang and Holtslag*, 1996]. A comparison of modeled PBL heights against radar observations at a tall tower in Wisconsin [*Yi et al.*, 2000] and aircraft tracer-derived observations (C. Gerbig et al., Constraining regional and continental fluxes of $CO_2$ with atmospheric observations over a continent: A receptor oriented analysis of the COBRA data concentration measurements, submitted to *Journal of Geophysical Research,* 2002, hereinafter referred to as Gerbig et al., submitted manuscript, 2002) shows that the new parameterization has significantly reduced the bias in the predicted PBL height (Figure 3).

[23] STILT adopted a unique way for transporting particles between vertical levels in order to satisfy the well-mixed criterion. It is important to note that most numerical treatments violate the well-mixed criterion in an environment with inhomogeneous turbulence, allowing initially well-mixed particles to accumulate in low-turbulence regions. A "drift correction" to address this problem was described by *Legg and Raupach* [1982], albeit it was later shown that this simple correction does not strictly satisfy the well-mixed criterion [*Rodean*, 1996; *Thomson*, 1984]. We initially adopted the drift correction and conducted tests to

verify adherence to the well-mixed criterion in the model. We turned off mean advection and prescribed time invariant vertical profiles of $T_{\mathrm{Lw}}$ (vertical) and $\sigma_w$, released particles evenly throughout the atmosphere, and monitored for signs of improper convergence. Particles remained well mixed for simple profiles with $T_{\mathrm{Lw}}$ and $\sigma_w$ varying linearly with height in the PBL. But when more complicated profiles of $T_{\mathrm{Lw}}$ and $\sigma_w$ were prescribed, particles accumulated in low-turbulence regions. *Thomson et al.* [1997] outlined a reflection/transmission scheme for Gaussian turbulence that preserves well-mixed distributions for particles moving across interfaces between step changes in turbulence parameters. We adopted this approach to satisfy the well-mixed criterion in the strongly inhomogeneous environment of the PBL, where the simple drift correction does not work. Vertical variations of $\sigma_w$ and $T_{\mathrm{Lw}}$ in STILT are approximated as discontinuous layers with constant values. A typical configuration in STILT has 17 layers below the altitude of 10 km, with denser coverage near the ground (five layers below 1 km). A particle with vertical velocity $w_i$ and approaching from below an interface between lower ($z_{i-}$) and upper ($z_{i+}$) layers is randomly transmitted with probability $\alpha$ and velocity $w_t$ or otherwise reflected with $-w_i$ [*Thomson et al.*, 1997], where

$$w_t = w_i \frac{\sigma_w(z_{i+})}{\sigma_w(z_{i-})}, \quad \alpha = \frac{\sigma_w(z_{i+})\rho(z_{i+})}{\sigma_w(z_{i-})\rho(z_{i-})}. \tag{9}$$

The same scheme is applied for a particle traveling downward and approaching the interface from above by reversing the lower and upper variables. The air density is represented as $\rho$; this term is added to the formulation of *Thomson et al.* [1997] because particles in STILT are treated as air parcels with constant mass, and changes in air density have to be taken into consideration in order to conserve the amount of mass each particle represents as it is transported vertically. Equation (9) ensures that initially well-mixed particles in an environment with Gaussian turbulence remain well mixed. Our tests for well mixedness showed no spurious accumulation for model run times of over 72 hours.

[24] STILT is currently driven by archived meteorological data from the Eta data assimilation system (EDAS) at the National Centers for Environmental Prediction [*Black*, 1994]. EDAS has a 32 km, 45 level grid and archives data in 3 hourly intervals with 80 km horizontal resolution and 22 vertical levels. Gridded variables in EDAS were interpolated in time and space to each particle location. Surface fluxes of momentum and sensible heat from EDAS were used in the parameterization of $T_{\mathrm{Lw}}$ and $\sigma_w$ [*Hanna*, 1982]. EDAS winds were directly used for the mean (nonstochastic) velocity $\bar{u}$. Meteorological variables were linearly interpolated between the 3 hourly intervals in EDAS and extracted at 1 min time steps. We implemented operator splitting, using finer time steps to properly resolve dispersion. Wind vectors from EDAS were passed onto the dispersion module to capture the interaction between windshear and vertical turbulence: particles are mixed by vertical turbulence to various heights, and the differing mean wind vectors at those heights advect the particles in differing directions and speeds. Particle locations were updated in the dispersion module at time steps set at 10% of the local $T_{\mathrm{Lw}}$ in the vertical profile to adequately resolve the decay in autocorrelation [*Wilson and Zhuang*, 1989]. $T_{\mathrm{Lw}}$ can vary from $\sim$1 s to hundreds of seconds, so the dispersion time steps can vary correspondingly from 0.1 s to tens of seconds.

## 5. Empirical Assessment of Time Reversibility

### 5.1. Comparison Between Forward and Backward Time Simulations

[25] An empirical test of time reversibility was conducted for the STILT model as an effective way to probe the model's deficiencies, revealing deviation from the well-mixed criterion through a comparison between the forward and backward time directions. The time reversibility tests also represent a practical, unique method to examine compliance with the well-mixed criterion in three dimensions; the alternative method of initializing particles well mixed throughout the entire atmosphere and tracking them to look for spurious convergence over time is not feasible, given the huge computational expense this entails. Furthermore, establishment of time reversibility enables the forward time results to be derived with backward time simulations, potentially allowing significant savings in computational effort.

[26] The test of time reversibility was conducted as follows. Time reversibility requires $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ derived from an ensemble of particle trajectories running either forward or backward in time to be the same. Values of $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ are derived from particle density (equation (4)) or, when integrated over a finite volume element, the number of particles in the volume element (equation (5)). Thus given that $N_{\mathrm{tot}}$ particles are released from the receptor location $\mathbf{x}_r = (x_r,y_r,z_r)$ at time 0 and transported backward in time for $t$ hours, and a three-dimensional box centered at $(x_i',y_i',z_i')$ encloses $n$ particles, a release of $N_{\mathrm{tot}}$ particles from $(x_i',y_i',z_i')$ at time $-t$ should also yield $n$ particles in a box with the same dimensions centered at $(x_r,y_r,z_r)$ after $+t$ hours (Figure 1), after taking into account differences in air density at $(x_r,y_r,z_r)$ and $(x_i',y_i',z_i')$. The dependence on air density follows from the fact that particles are treated as air parcels with mass, so the number of particles found in a volume at the source and receptor will depend on the local air density at the receptor and source locations.

[27] For our test we set the receptor location $(x_r,y_r,z_r)$ to the WLEF television station tall tower monitoring station in Wisconsin [*Bakwin et al.*, 1996]. Particle trajectories were simulated using Nested Grid Model (NGM) meteorological fields [*Hoke et al.*, 1989], due to the 180 km NGM fields' coarser resolution relative to EDAS and smaller associated computational expense. Eight starting times were randomly selected in 1996, when 15,000 particles were released from a box centered at 250 m above ground level (AGL) and transported 48 hours backward in time. The source region defined by the particle locations at $-48$ hours was divided into boxes centered vertically at 50 m AGL, and the number of particles found in each box was counted. The box size was chosen to have a vertical dimension of 100 m, and horizontal dimensions were selected to roughly divide the source region into 100 boxes. The horizontal dimensions ranged from 0.5° longitude by 0.2° latitude to 5° longitude by 1° latitude between the eight starting times. We released

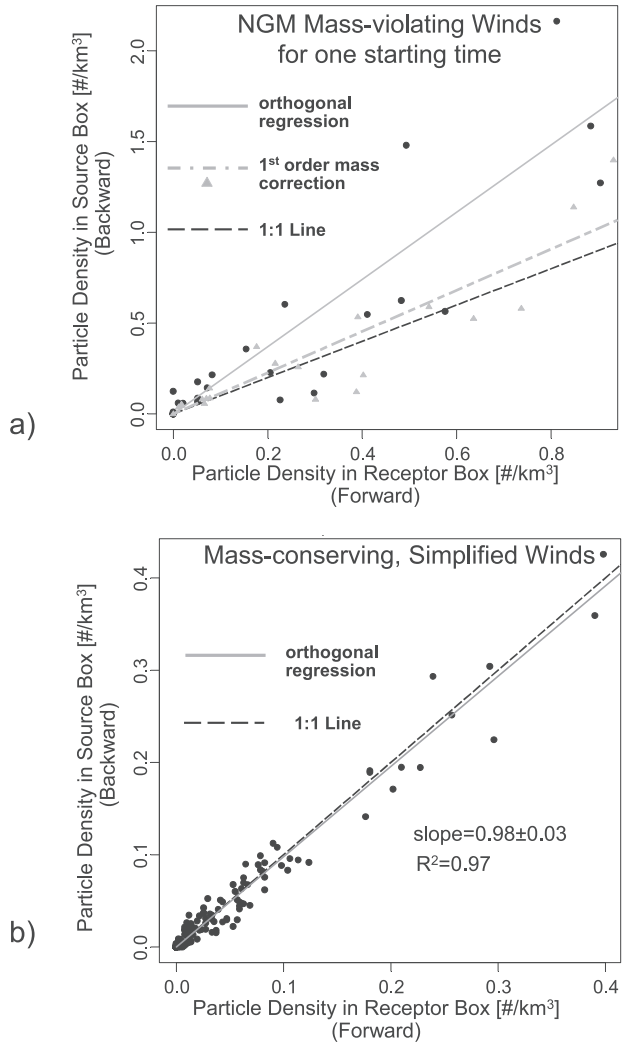


**Figure 4.** (a) Comparison of particle densities from 48 hour backward and forward time model runs starting on 10 April 1996 at 0800 UT. The triangles denote particle densities after application of a first-order mass correction, in which particles traveling through regions of mass creation receive heavier weight while particles in regions of mass destruction receive lighter weight (see text). (b) Comparison of forward and backward time model runs when a simplified, mass-conserving wind field is used.

15,000 particles from a subset (25%) of the boxes, transported them 48 hours forward in time, and tallied the particles ending up in the receptor box. The particle density in the source box was then compared with the corresponding number in the receptor box. Subgrid level heterogeneities in particle density not resolved by the source and receptor boxes were removed by distributing the starting locations of particles uniformly throughout the boxes for both the forward and backward simulations.

[28] Comparison for a single starting time, at 0800 UT on 10 April, is shown in Figure 4a as an example of the results. The backward time particle densities were correlated with their forward time counterparts ($R^2$ = 0.80), but the comparison exhibited large amounts of scatter, with a significant bias toward higher particle densities in the backward simulations; the slope given by a weighted orthogonal distance regression was 1.85 ± 0.11. The weights (≡1/error) in the regression were assigned by treating errors in particle numbers to follow Poisson statistics, as confirmed by bootstrap analysis. Comparisons for all eight starting times are shown in Table 1. The regression slopes were statistically different from 1 for almost all the days examined, suggesting temporal asymmetry in the NGM particle simulations.

### 5.2. Effect of Mass Violation

[29] The asymmetry between backward and forward simulations in Figure 4a can be traced to mass violation in the driving meteorology, implying significant errors in source-receptor influence functions for both forward and backward models. Mean wind trajectories simulated by STILT are reversible with negligible errors (<1 km horizontal,<2 m vertical over 48 hours), suggesting that the advection scheme on its own is not the source of the asymmetry. Furthermore, when mean advection was switched off, the turbulence parameterization by itself was reversible, satisfying the well-mixed criterion. These results strongly suggest that the problem likely resides in the driving winds instead of the numerical representation of particle transport in the model. Asymmetry under time reversal due to mass imbalances can be introduced via spurious dilution and concentration of particles, as discussed earlier (Figure 2).

[30] The wind fields from numerical weather prediction centers could violate mass conservation due to coordinate transformations during the postprocessing procedure or due to the data assimilation process itself [*Byun*, 1999]. *Trenberth* [1991] has shown that large mass imbalances can arise from the way in which model variables are postprocessed and archived, e.g., mass imbalances can arise from interpolation of variables to individual levels rather than averages over entire layers. Degradation of spatial resolution during the postprocessing and archiving process can clearly introduce mass violation, especially when grid points are simply omitted instead of used in an averaging scheme over the coarser grid. Conversions between vertical coordinate systems can introduce additional errors [*Hoerling and Sanford*, 1993].

[31] The NGM meteorological files used in the time reversibility tests were degraded from a spatial resolution

**Table 1.** Results of the Empirical Tests for Time Reversibility That Compare 48 Hour Backward and Forward Time Simulations of the STILT Model (See Text)[a]

| Day | Hour, UT | Regression Slope | Regression Slope (First-Order Mass Correction) |
|---|---|---|---|
| 28 Jan. | 0900 | 1.01 ± 0.06 | 0.96 ± 0.06 |
| 4 March | 2000 | 0.68 ± 0.05 | 0.80 ± 0.07 |
| 7 March | 1800 | 1.03 ± 0.07 | 0.98 ± 0.07 |
| 1 April | 1700 | 0.68 ± 0.07 | 0.67 ± 0.07 |
| 10 April | 0800 | 1.85 ± 0.11 | 1.13 ± 0.07 |
| 20 Feb. | 0100 | 0.18 ± 0.03 | 0.52 ± 0.03 |
| 25 March | 0700 | 0.97 ± 0.07 | 0.96 ± 0.06 |
| 10 May | 2100 | 1.17 ± 0.07 | 0.91 ± 0.06 |

[a]The slope (backward/forward) is derived from an orthogonal distance regression. The last column refers to the regression slope after applying a first-order correction for mass violation in the NGM fields that assigns weights to particles depending the cumulative mass violation that they have experienced during their transport history.

of 90–180 km during the archival procedure. We quantified mass violation in the NGM fields by directly calculating imbalances in the continuity equation using wind vectors and pressure changes from the NGM fields. Horizontal and vertical mass divergence terms were calculated from differences in mass fluxes between adjacent grid points. Large values of mass violation were observed, with as much as 15% of mass in a gridcell being created or destroyed per hour at the lowest model level and 2% at higher model levels.

[32] A first-order correction for the NGM particle simulations was carried out by tallying the cumulative mass violation experienced by each particle during its transport history. We then allowed particles traveling through regions of mass creation in the meteorological field to become correspondingly "heavier"; conversely, particles in regions of mass destruction become "lighter." Each particle is assigned a weight according to its fraction of the total mass of the particle ensemble at a particular time step. Mass creation giving rise to particle dilution and an underestimation of particle density would be partially corrected by the greater weight assigned to the resulting heavier particles, while particle concentration associated with mass destruction would be corrected by deemphasizing the lighter particles (Figure 2). This method provides only a first-order correction effective only if the particle ensembles follow similar paths traveling to the receptor and the source; errors arising from mass violation can be amplified due to heterogeneity in the three-dimensional wind field as particle dilution/concentration forces particles to take forward time paths that are different from their backward time counterparts.

[33] Application of the mass correction to the comparison between forward and backward time simulations generally resulted in a decrease in the temporal asymmetry (Figure 4a and Table 1), with the regression slope shifting closer to 1.0 for five out of the eight starting times. The improvement in correspondence resulting from a correction incorporating mass considerations supports the view that mass violation in the driving wind field introduces asymmetries into the particle results.

[34] As further investigation into the effects of mass violation on particle simulations, we constructed a simplified, mass-conserving atmosphere by prescribing winds with simple linear dependence on altitude and no horizontal divergence ($\partial u/\partial x = \partial v/\partial y = 0$). The vertical mean wind $w$ was set to 0 everywhere. A sinusoidal temporal dependence of the $v$ wind was prescribed to cause horizontal dispersion of particles. Both air density and ground height were set to constant values throughout the model domain. Turbulence parameters were taken directly from the NGM meteorological fields without simplification, retaining the diurnal variations in PBL height and vertical mixing. When the same comparison between forward and backward time simulations was conducted using the mass conserving wind field described above, the bias was no longer observed (slope = 0.98 ± 0.03), with $R^2$= 0.97 (Figure 4b). This suggests that particle transport as modeled by STILT is time reversible, provided the particles are transported by mass conserving wind fields.

[35] We then examined qualitatively the effect of large mass violation on time reversibility in the simplified atmosphere described above by prescribing a wind field with mass violation of 15% $h^{-1}$ below 1800 m AGL by setting vertical divergence to a nonzero value ($\partial w/\partial z$ = −15% $h^{-1}$). Thus particles traveling forward in time below 1800 m experience mass destruction, as the linear increase in subsidence velocity is not balanced by horizontal divergence. Vertical velocity at higher levels was set to zero, with the velocity linearly decaying to zero between 1800 m and the next model level at ~2200 m. This is a region of significant mass creation (>100% $h^{-1}$) for forward time particles.

[36] The effect of this extensive deviation from mass conservation is clearly manifested in backward and forward time particle runs showing particle number as a function of altitude and time (Figures 5a and 5b). The backward time simulation (Figure 5a) shows dilution of particles in the lower atmosphere, where mass creation occurs, and accumulation of particles in a thin layer near 2000 m, where extensive mass destruction takes place. The strong effect of mass violating winds is only partially counteracted during hours of more extensive vertical mixing, coinciding with higher PBL heights. In contrast, the forward time simulation (Figure 5b) shows particles being trapped near the ground, where mass destruction of 15% $h^{-1}$ takes place. The dependence of mass destruction and creation on the direction of time has led to asymmetries in particle simulations (Figure 2). Consider a receptor and source that are both located near the ground; few particles traveling backward would be found at the source, while many more particles traveling forward would reach the receptor.

[37] The combination of (1) lack of a significant bias in the forward-backward comparison in the simplified, mass conserving atmosphere and (2) the large effect of mass violation further supports the view that the observed time asymmetry in Figure 4a and Table 1 is mostly due to mass violation in the NGM winds. Evidently, inconsistencies in archived wind fields leading to mass violation must be addressed to obtain accurate results using STILT or any other offline transport model.

## 6. Application of STILT to Atmospheric Observations: COBRA Experiment

### 6.1. Need for a Near-Field Tool to Interpret $CO_2$ Observations

[38] We illustrate the usefulness of applying STILT for interpreting aircraft observations of $CO_2$. Terrestrial sources and sinks of $CO_2$ are widely distributed at the surface, where the biosphere acts as a sink during daytime and a source during nighttime. The direction of surface fluxes is correlated in time with the PBL depth, rendering interpretation of $CO_2$ observations in the PBL difficult without adequate representation of PBL dynamics [*Denning et al.*, 1996]. The $CO_2$ signal in the lower atmosphere shows significant variability even down to length scales of 10 km (Gerbig et al., submitted manuscript, 2002), suggesting that gradients due to variable sources and sinks of $CO_2$ have not been averaged out by dispersion and mixing. These characteristics call for a tool like STILT, which can interpolate wind fields down to the subgrid scale and adequately resolve near-field transport in the PBL.

[39] The data analyzed here derive from test flights of the COBRA study. COBRA was designed to address the current

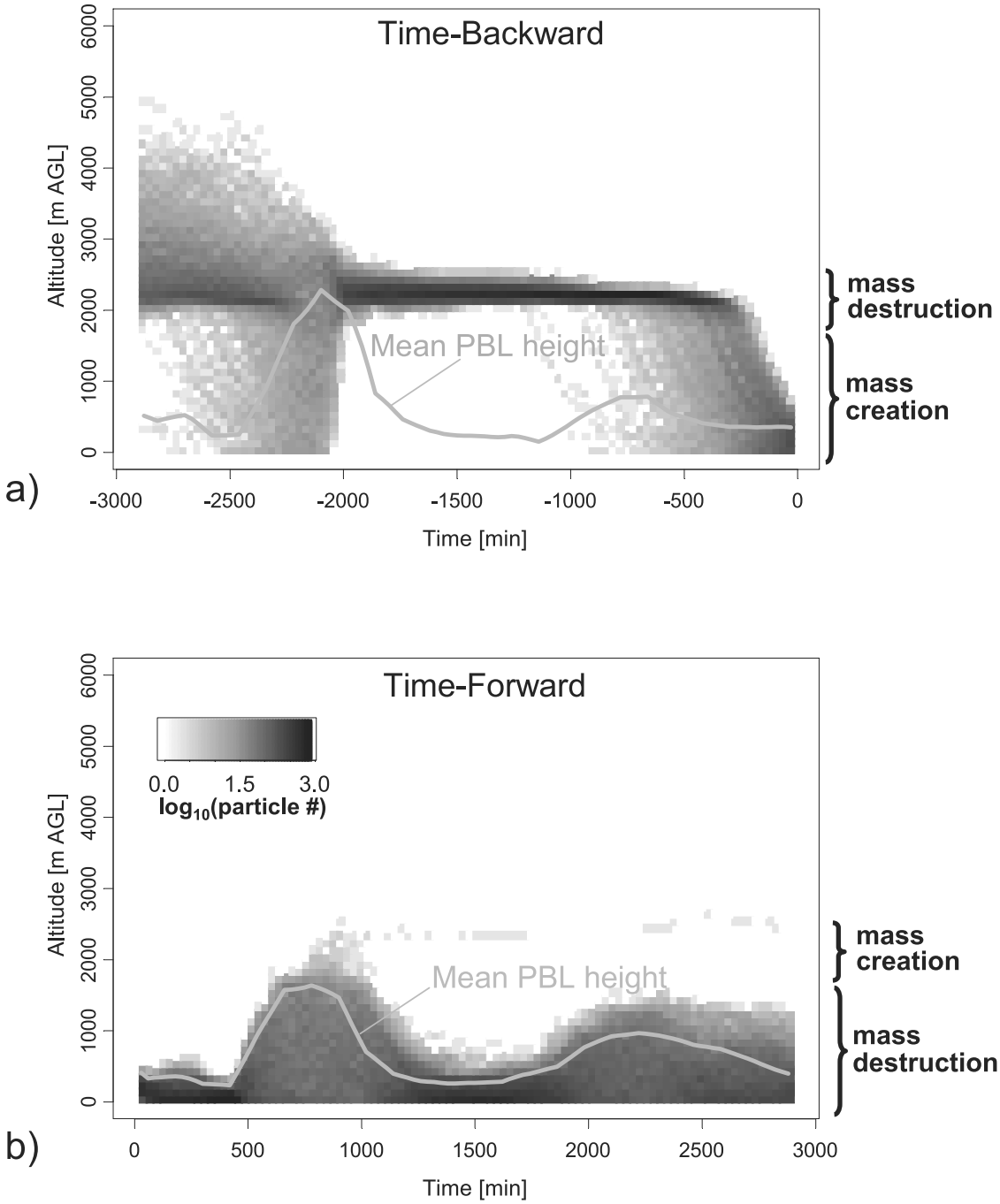


**Figure 5.** Particle runs driven with mass-violating wind fields in a simplified atmosphere. Particle simulations (a) backward and (b) forward in time, driven by winds with 15% mass violation below 1800 m above ground level and >100% between 1800 and 2200 m.

gap in knowledge of terrestrial carbon fluxes at the regional to continental scales [*Stephens et al.*, 1999]. We ran STILT in the backward time mode to interpret $CO_2$ observations during the 1999 COBRA test flights, as part of which the University of North Dakota Cessna Citation aircraft acquired a time series of vertical profiles in the vicinity of the WLEF tall tower on 8 June (Figure 6). The airborne $CO_2$ instrument used in COBRA is a nondispersive infrared gas analyzer based on the design of the Harvard University ER-2 $CO_2$ analyzer [*Andrews et al.*, 1999; *Boering et al.*, 1994]. In-flight and laboratory calibrations indicate that the uncertainty of the $CO_2$ observations during COBRA was ±0.25 ppm [*Daube et al.*, 2002]. A distinct decrease of $CO_2$ near WLEF over the day can be seen in Figure 6, presumably due to biospheric uptake.

[40] Similar observations have been used to calculate regional $CO_2$ uptake with a boundary layer budget method by assuming horizontal advection of tracers into the atmospheric column to be small [*Denmead et al.*, 1996; *Kuck et al.*, 2000; *Levy et al.*, 1999; *Lloyd et al.*, 2001], essentially assuming that upstream regions influencing the observations are identical in both the morning and the afternoon,

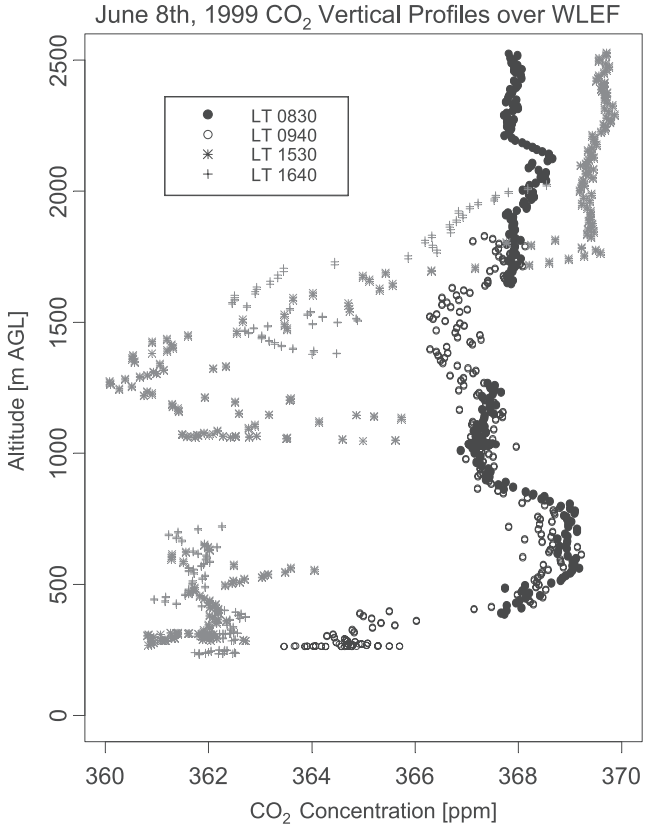


**Figure 6.** $CO_2$ vertical profiles over WLEF at different local times on 8 June 1999 during the COBRA test flight series. The scatter in data points for 1530 LT near 1500 m can be attributed to eddies that entrained air from above the mixed layer, as seen in the similarities in tracers such as CO, $H_2O$, and θ with the air above the mixed layer. Gaps in the vertical profiles resulted from in-flight calibrations.

with the tracer decrease reflecting daytime uptake by that same upstream region. The boundary layer budget can be understood from a receptor-oriented perspective by integrating equation (3) vertically to the top of a column and treating $C$ as the vertically integrated column tracer amount. The difference in column tracer amount at the receptor ($\Delta C$) during two times ($t_{r1}$, $t_{r2}$) is then

$$\begin{aligned}\Delta C &= C(\mathbf{x}_r, t_{r2}) - C(\mathbf{x}_r, t_{r1}) \\ &= \left[\int_{t_0}^{t_{r2}} dt \int_V d^3x I(\mathbf{x}_r, t_{r2} \mid \mathbf{x}, t) S(\mathbf{x}, t) \int_V d^3x I(\mathbf{x}_r, t_{r2} \mid \mathbf{x}, t_0) C(\mathbf{x}, t_0)\right] \\ &- \left[\int_{t_0}^{t_{r1}} dt \int_V d^3x I(\mathbf{x}_r, t_{r1} \mid \mathbf{x}, t) S(\mathbf{x}, t) + \int_V d^3x I(\mathbf{x}_r, t_{r1} \mid \mathbf{x}, t_0) C(\mathbf{x}, t_0)\right].\end{aligned} \tag{10}$$

By making the assumption that $I(\mathbf{x}_r,t_{r2} \mid \mathbf{x},t) = I(\mathbf{x}_r,t_{r1} \mid \mathbf{x},t) \equiv I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$, which corresponds to the assumption that footprints are similar at times $t_{r1}$ and $t_{r2}$ and that horizontal advection is negligible from a Eulerian perspective, terms involving the initial field $C(\mathbf{x},t_0)$ cancel, and $\Delta C$ can be linked directly to $\ll S \gg$ the spatially and temporally averaged source emission, in a simple framework:

$$\begin{aligned}\Delta C &= \int_{t_0}^{t_{r2}} dt \int_V d^3x I(\mathbf{x}_r, t_r \mid \mathbf{x}, t) S(\mathbf{x}, t) \\ &\quad - \int_{t_0}^{t_{r1}} dt \int_V d^3x I(\mathbf{x}_r, t_r \mid \mathbf{x}, t) S(\mathbf{x}, t) \\ &= \int_{t_{r1}}^{t_{r2}} dt \int_V d^3x I(\mathbf{x}_r, t_r \mid \mathbf{x}, t) S(\mathbf{x}, t) = (t_{r2} - t_{r1}) \ll S \gg .\end{aligned} \tag{11}$$

The assumption of $I(\mathbf{x}_r,t_{r2} \mid \mathbf{x},t) = I(\mathbf{x}_r,t_{r1} \mid \mathbf{x},t)$ must be critically assessed. Heterogeneities in land cover may be associated with large differences in $S(\mathbf{x},t)$ that translate into concentration differences at the receptor. This situation is particularly important in the near field of observations, where sources/sinks have strong influence on tracer concentrations, and where atmospheric dispersion and mixing have not yet averaged over multiple vegetation types. Whether air traveled over corn fields or the forest during the previous 24 hours prior to arriving at WLEF, for instance, is important for interpreting the $CO_2$ observations. We use STILT to generate particle distributions that can exhibit detailed structures in the near field and quantitatively resolve $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$. By linking $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ to the footprint of an receptor observation, the sensitivity of an observation to upstream surface fluxes (equation (8)), the relative importance of different vegetation types upstream of the receptor can be assessed, providing the capability to examine whether horizontal advection can be neglected.

### 6.2. Derivation of the Footprint

[41] Particle trajectories were used to generate the footprint elements $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$, which provide spatially resolved, quantitative information linking upstream surface fluxes with the receptor. We released particles at five heights (50, 500, 1000, 1500, and 2000 m AGL; $N_{tot}$ = 8464, 8000, 7696, 7308, and 5604 particles, respectively) over WLEF during the hours when the aircraft was conducting vertical profiles, and particles were transported back in time for 24 hours using the EDAS meteorological field. $N_{tot}$ at each release height was scaled according to air density such that the $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ generated by all the particles is representative for the entire vertical profile. The $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$ was quantified from particle locations using equation (8). The above first-order mass correction was applied in an attempt to minimize artifacts in the simulations arising from mass violation in the EDAS fields. Simulated footprints from STILT were found to be insensitive to the exact value of column height $h$ between 10 and 100% of the PBL height (Gerbig et al., submitted manuscript, 2002); we simply equated $h$ to the PBL height here.

[42] The footprints generated by STILT for 0800 LT and 1600 LT at −6, −12, and −24 hours are shown in logarithmic scale in Figure 7, with surface fluxes in regions indicated by dark squares causing a greater concentration change at the WLEF receptor. These results show that the regions influencing the WLEF receptor changed between the morning and afternoon. Westerly/northwesterly winds at

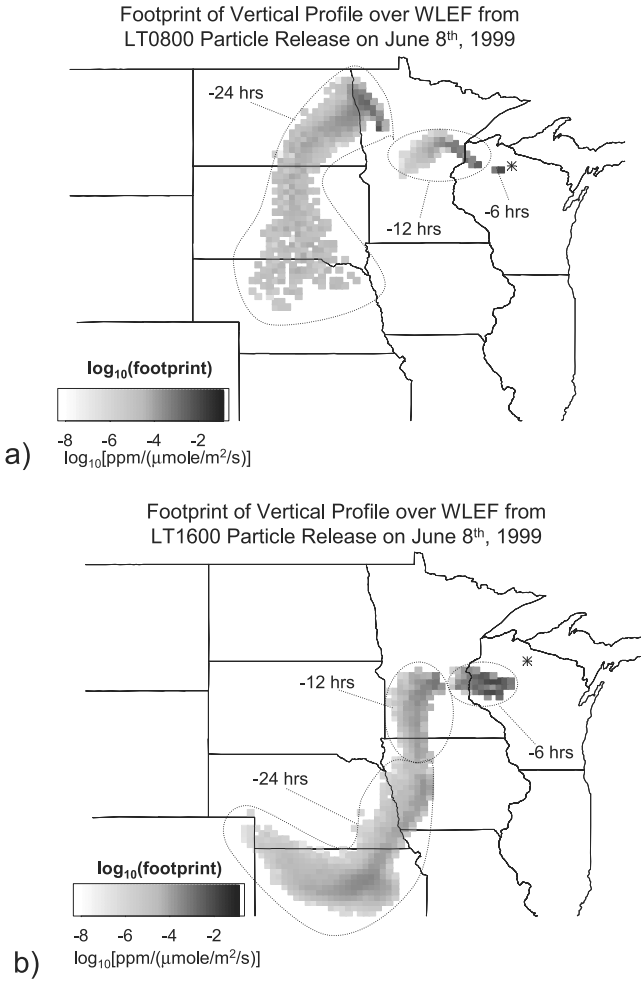


**Figure 7.** Footprints derived from particle locations generated by STILT model at −6, −12, and −14 hours upstream from WLEF (asterisk) on 8 June 1999 at (a) 0800 LT and (b) 1600 LT. Particles were released over altitudes covering the aircraft vertical profile observations shown in Figure 6, such that the footprints are applicable for the entire vertical profile. The greyscale shows the logarithm (base 10) of the footprint in each 1/6° latitude by 1/4° longitude gridcell. Darker areas denote regions where a unit surface flux leads to a greater change in concentration at WLEF. A clear shift in transport history of air arriving at WLEF can be seen over the day between 0800 LT and 1600 LT.

0800 LT shifted to southwesterly winds by 1600 LT. Possible evidence of a wind shift can also be seen in the free tropospheric value of $CO_2$, which increased from 368 to 369 ppm during the day (Figure 5). Evidently, advective influences cannot be neglected in this example, and changes in $CO_2$ do not simply reflect contributions from surface sources and sinks.

[43] We examined the consequence of the wind shift on relative influences from different vegetation types on the receptor concentration. The footprint simulated by STILT was superimposed on the 1 km International Geosphere Biosphere Program land classification grid [*Belward et al.*, 1999], regridded to 1/6° latitude by 1/4° longitude. Footprints from the four dominant vegetation classes at each hour of transport during the day prior to arrival at WLEF (at 0800 LT and 1600 LT on 8 June) are plotted in Figure 8. The rapid decay in footprint strength during the first hours moving backward from 0800 LT was due to strong vertical winds, which lofted particles above the shallow nocturnal PBL. The shift in winds seen in Figure 7 translated into differences in vegetation footprints during the 24 hours before arrival at WLEF. Air arriving at 1600 LT was heavily influenced by mixed forests and the cropland/natural vegetation mosaic during the day, while daytime mixed forest footprint was not as strong for 0800 LT. Cropland footprint was high during the entire night for the 1600 LT air as it was transported back in time, whereas cropland footprint for the 0800 LT air did not increase until the previous day's afternoon.

[44] Accurate footprints $f(\mathbf{x}_r, t_r \mid x_i, y_j, t_m)$ generated by a tool like STILT are essential for interpreting atmospheric observations and determining the applicability of an one-dimensional budget approach. The footprints link fluxes from upstream sources/sinks with the tracer observations. Instances with changes in footprint (Figure 7) can be identified when a simple framework like equation (11) to derive fluxes from directly differencing tracer amounts cannot be applied. Footprint change, correlated with spatial heterogeneity in vegetation cover, translated into different vegetation types influencing the receptor (Figure 8). Because characteristics of nighttime release and daytime uptake of carbon may differ significantly between vegetation types, the shift in vegetation footprint on this day means that a well-defined $\ll S \gg$ cannot be calculated from simple differences in tracer profiles. Furthermore, the change in wind direction would lead to different values in the initial field $C(\mathbf{x},t_0)$ to be advected to the receptor location, as suggested by the observed increase in the free tropospheric value of $CO_2$ (Figure 6). It follows that one cannot assume that terms involving the initial field $C(\mathbf{x},t_0)$ cancel, and the simplified framework (equation (11)) does not apply.

### 6.3. Alternative Approach to Estimating Fluxes: "Influence-Following" Experiments

[45] Alternative measurement approaches from a simple one-dimensional boundary layer budget may be necessary to derive fluxes during times when the assumption of a constant $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ between two sampling times does not hold. *Chou et al.* [2002] adopted a pseudo-Lagrangian framework to derive fluxes of $CO_2$ from an atmospheric budget, made possible by the availability of numerous vertical profiles over multiple weeks during the ABLE-2B mission and the vast extent of the Amazonian rain forest. Furthermore, more flexible flight planning would enable an experiment which samples $C(\mathbf{x}_r,t_r)$ and $C(\mathbf{x},t_0)$ that, combined with knowledge of $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$, imposes direct constraints on the spatially integrated flux between $t_0$ and $t_r$ (equation (3)). $C(\mathbf{x},t_0)$ refers to the initial tracer field over a large domain, but only the subdomain where $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ is nonzero, i.e., where upstream air parcels influencing the receptor are found, needs to be characterized. An observational strategy that tracks air parcels in the atmosphere, and particles in STILT, can characterize $C(\mathbf{x},t_0)$. Note that this is not a simple "air mass following" experiment, because

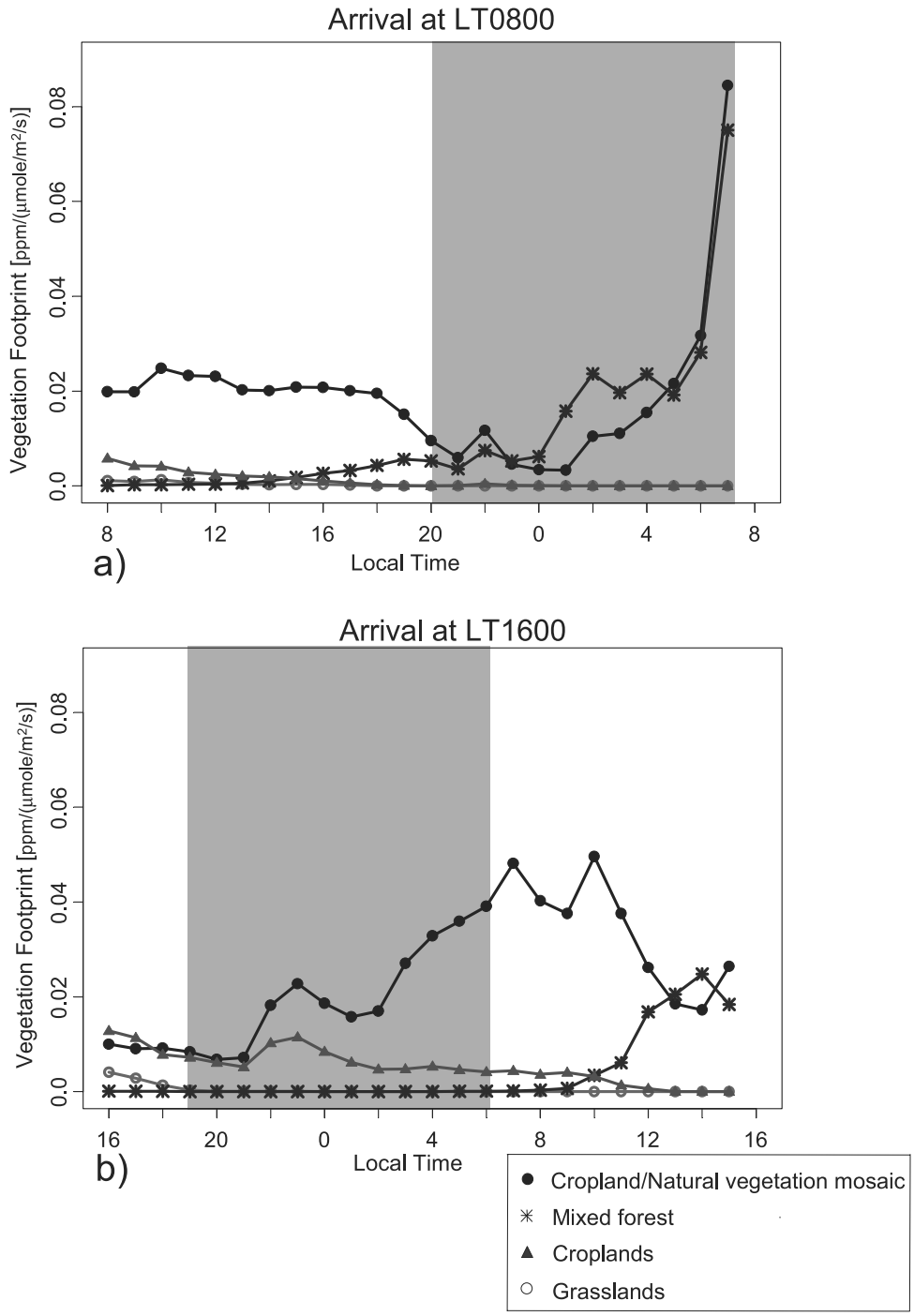


**Figure 8.** Time series of footprints due to the four main vegetation classes during the 24 hours prior to arrival on 8 June 1999 at (a) 0800 LT and (b) 1600 LT. Shaded portions denote nighttime periods. At each time step, particle locations were mapped onto the IGBP 1 km resolution vegetation data, regridded to 1/6° latitude by 1/4° longitude resolution. At the particle locations, footprints from different vegetation classes were quantified according to equation (8), with the time interval $\tau$ set to 1 hour. An unit flux of 1 μmol $m^{-2}$ $s^{-1}$ by a particular vegetation class persisting over an hour would lead to the change in $CO_2$ concentration at WLEF specified by the vegetation footprint. See color version of this figure at back of this issue.

windshear and dispersion can cause air parcels arriving at the receptor to originate from different air masses, but an ''influence-following'' experiment to capture upstream influences. Implementing an influence-following experiment requires the upstream influences $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ to be generated ahead of time in order to plan flight tracks that cover all of the relevant domain. STILT, driven with forecasted meteorology from the Eta model, served as an

operational flight planning tool for determining $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ in the COBRA flights of August 2000. Flights were planned to sample the locations marked out by particles modeled with STILT. A comprehensive analysis of the COBRA influence-following experiments will be presented in a later publication (J. C. Lin et al., manuscript in preparation, 2003).

## 7. Information Gain From Application of Near-Field Tool

[46] Since STILT interpolates winds down to the subgrid scale and tracks air parcels upwind to reveal influences, it is expected to provide information more detailed and realistic than a Eulerian approach, where the influence on a receptor comes from upwind areas which may only be poorly resolved by Eulerian grids. The gain in information is particularly important in the near field, where surface influence on the receptor is strong (Figure 8) and where influence areas can show detailed structures at the subgrid scale (Figure 7).

[47] We draw ideas from information theory to provide a quantitative comparison of information about influence areas provided by the time-reversed Lagrangian particle approach versus the conventional Eulerian gridcell approach. Information can be expressed mathematically as a reduction in uncertainty [*Brillouin*, 1956], where uncertainty is quantified by the Shannon entropy $H$ [*Shannon and Weaver*, 1963] as follows:

$$H = \sum_i p_i \log_2 \left(\frac{1}{p_i}\right). \tag{12}$$

Here $p_i$ denotes the probabilities of various outcomes $i$ of a random variable. The base 2 in the logarithmic function treats the unit of information as the "bit," such that information can be thought of as the minimum number of binary digits required to encode it [*Renyi*, 1984].

[48] Applied to influence areas, $H$ quantifies the uncertainty about the potential emission source and its location in the area, and $p_i$ represents the probability of the emission source coming from a region $i$. Atmospheric dispersion and mixing cause increases in $H$ (or losses in information), reflected in the spread of the influence region over larger and larger regions. Given a specific size of an influence region, the value of $H$ is largest when $p_i$ is uniform, i.e., when influence is spread out evenly over the entire influence area, there is greater uncertainty about the source than if influence were concentrated in small subregions. Eulerian simulations spatially resolve influence areas at the grid scale. Lagrangian particles can resolve influences at the subgrid scale, so greater uncertainty (less information) about the emission source is associated with the Eulerian result. Note that the information gain calculated here is the reduction in uncertainty in the location of the source as given by the influence region; quantifying $I_{\text{gain}}$ for "specific tracers" will require a more involved calculation that takes into account not just the influence area, but also the spatial variability of sources/sinks in the influence area that give rise to tracer variations.

[49] We define the information gain ($I_{\text{gain}}$) from adopting a Lagrangian particle approach as the reduction in Shannon entropy from Eulerian model results ($H_E$) to the particle results ($H_L$):

$$I_{\text{gain}} = H_E - H_L. \tag{13}$$

[50] The particles provide a straightforward calculation of $H$, since the fraction of particles found in a particular gridcell $i$ over a time interval, as expressed by equation (5), can be interpreted as $p_i$, the probability of getting influence from gridcell $i$. The values of $p_i$ for quantifying $H_E$ are calculated by evenly distributing particles found in a Eulerian gridcell over the entire 80 km $\times$ 80 km (EDAS resolution) surface, as the Eulerian approach does not resolve influence areas at the subgrid scale. In the case of $H_L$ the idea of "gridcells" is not as clearly defined, but should reflect the limit in resolution of the Lagrangian approach, i.e., within the chosen size of the gridcell particle positions are uncertain, and no further information could be derived from their positions. Uncertainties in the Lagrangian approach arise from uncertainties in particle position associated with interpolating wind fields to derive the subgrid scale winds in the particle simulations. *Stohl et al.* [1995] have found the interpolation error in horizontal wind ($\delta u$) to be $\sim$0.4 m s$^{-1}$ when the magnitude $u$ of the wind is 10 m s$^{-1}$; this error would cause an error $\varepsilon$ in particle position over its travel across the gridcell, given by $\varepsilon = [\Delta x/(u \pm \delta u)]\delta u$, where $\Delta x$ is the resolution of the wind fields. The value of $\varepsilon$ = 3.3 km is for the 80 km resolution EDAS data. We chose a gridcell of 20 km $\times$ 20 km, an order of magnitude larger, for $H_L$ as an upper bound on uncertainties in particle position associated with interpolating wind fields to derive a conservative estimate (lower bound) of information gain.

[51] The information gain provided by a particle approach at different hours backward in time from the receptor is shown in Figure 9. The receptor is chosen at WLEF, and particles were emitted at 1600 LT on 8 June 1999. Possible overestimation of information gain due to patchy coverage by a limited number of particles was ruled out by a sensitivity test doubling the particle number, which yielded the same results as in Figure 9. $I_{\text{gain}}$ is almost as large as 3 bits during the first few hours of transport back from the receptor, when the influence region is still small relative to the 80 km resolution Eulerian gridcell. As dispersion increases the size of the influence region, the Eulerian gridcells begin to resolve the influence region, leading to a decrease in $I_{\text{gain}}$. $I_{\text{gain}}$ stabilizes but remains at a nonzero value of $\sim$0.3 bits even over longer times. We can see that the information gain is large during the first $\sim$12 hours of transport backward from the receptor. The initial hours backward from the receptor are precisely the times when the additional information provided by a particle approach is important, due to the strong surface influence (Figure 8).

[52] The above estimate does not account for additional considerations that would further increase $I_{\text{gain}}$. The vertical resolution of influence was assumed to remain the same in both the Eulerian and Lagrangian approaches, whereas Lagrangian particle distributions can resolve influence fields at higher vertical resolution than the Eulerian grid levels. Moreover, numerical diffusion associated with the Eulerian approach [*Odman*, 1997] is disregarded when calculation of $H_E$ is derived from distributing particles

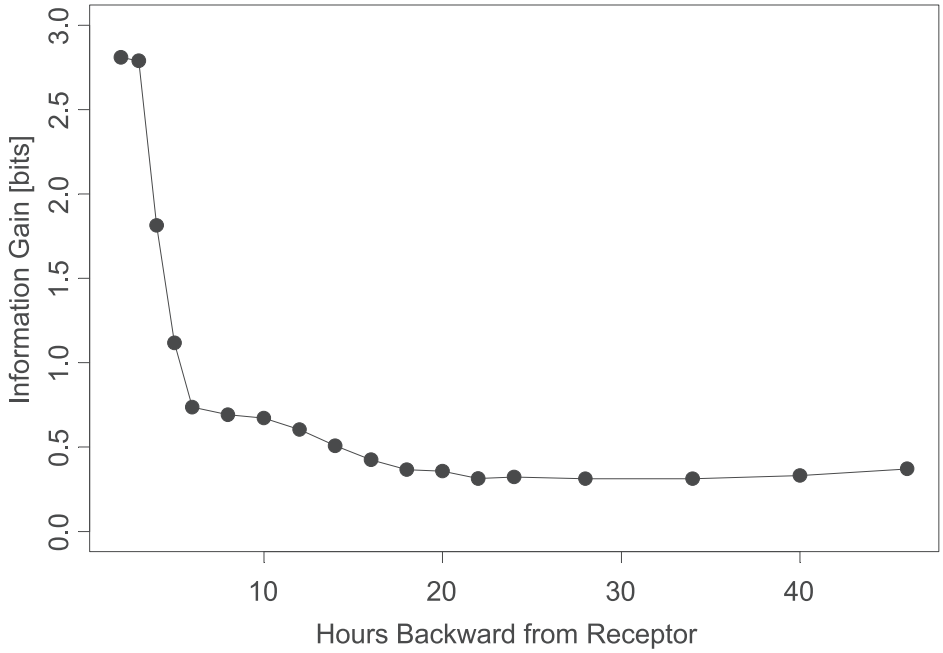


**Figure 9.** Information gain from Lagrangian particle approach over Eulerian gridcell approach as a function of hours backward from the receptor. The receptor is chosen at WLEF, and particles were emitted at 1600 LT on 8 June 1999.

evenly throughout a gridcell at each time step. The Eulerian approach dilutes tracers throughout a gridcell, and in subsequent time steps the tracer would be further dispersed to neighboring gridcells. The net effect is an atmosphere that is more diffusive than simulated by the particle approach, resulting in a larger value for $H_E$ and a higher $I_{\text{gain}}$ than shown in Figure 8.

## 8. Conclusions

[53] Time-reversed Lagrangian particle dispersion modeling is an effective way to determine the influence functions of surface sources on atmospheric trace gas concentrations observed at ground stations or by aircraft. Lagrangian particle approaches are particularly appropriate for simulation of near-field influence on concentrations by resolving subgrid scale transport while capturing the essential physics of turbulent transport in the PBL. The particle ensemble enables straightforward adoption of a receptor-oriented framework, providing upstream influences $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ on observations at the receptor (equations (3)–(5)). Particle simulations backward in time from the receptor serve as a particularly efficient way to derive the influence function $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$.

[54] But in order to yield realistic simulations of $I(\mathbf{x}_r,t_r \mid \mathbf{x},t)$ particle models have to satisfy physical criteria including the well-mixed criterion, resolution of the decay in autocorrelation, representation of close coupling between windshear and vertical turbulence, and consistent treatment of particles as air parcels with equal mass.

[55] STILT is a near-field tool designed to satisfy the above physical criteria and simulate turbulent dispersion and boundary layer transport on diurnal timescales, yielding upstream influences that quantitatively link observations with fluxes. The backward time simulation in STILT has been demonstrated to provide results consistent with its forward time counterpart. Discrepancies can be traced to mass violation in the driving meteorological fields, which can be attributed to the assimilation process or the post-processing process involving coordinate transformations. Future work will address the problem of mass imbalance in STILT by using winds on the native grid of the meteorological model, preserving the original spatial resolution and vertical coordinate system.

[56] We have shown that a tool like STILT can provide upstream influences in the near field that can be crucial for interpreting in situ observations. An information theoretic calculation for a case study has shown that the particle approach can yield significant gain in information about the location of the influence region over a standard Eulerian gridcell approach at least during the 12 hours before arrival at the receptor, when the observation is strongly influenced by surface fluxes and when near-field influence is important for interpreting the observations (Figure 8). We note that the exact information gain from a particle approach in specific situations depends on numerous factors not explored in Figure 9. The error due to interpolation, necessary in the Lagrangian approach, is expected to vary between meteorological situations, being larger and reducing the information gain during times of convective activity or in locations with complicated mesoscale circulations. The gain in information about tracer fluxes (rather than just the influence region as shown in Figure 9) depends upon the specific source/sink spatial distribution: the distance over which tracer fluxes are spatially correlated (e.g., between similar vegetation types for biospheric $CO_2$ fluxes) determines whether Eulerian gridcells can spatially resolve the contributions of tracer fluxes to receptor observations.

[57] A near-field tool like STILT has the potential to complement other existing modeling approaches. First, STILT can be coupled with large-scale models to link tracer observations at point locations over land with the global tracer field (''far field''). An illustration of the linkage between point observations of $CO_2$ and an oceanic boundary condition $C(\mathbf{x},t_0)$ to elucidate carbon fluxes can be found in the work of Gerbig et al. (submitted manuscript, 2002), who used STILT to link continental $CO_2$ observations with a climatological, oceanic $CO_2$ boundary condition generated by propagating surface observations [*Conway et al.*, 1994] to fit airborne observations higher in the atmosphere. Global tracer transport models make use of data at the global scale, providing an alternate means to generate the crucial far-field information $C(\mathbf{x},t_0)$ that can be coupled with the information provided by a near-field tool like STILT. Second, STILT complements the ''adjoint approach,'' which formally derives adjoint equations from Eulerian transport models [*Giering*, 1999; *Pudykiewicz*, 1998] that provide the sensitivity of model output (e.g., concentration) to input variables (e.g., fluxes) at the grid-scale [*Errico*, 1997]. Generation of the adjoint model source code can be a laborious and time-consuming process, which has to be implemented for each separate transport model. STILT can be driven off-line using archived winds, providing a rapid and flexible alternative. Even if the adjoint model is available, limitations due to finite-sized gridcells can result in inadequate resolution of the near-field influence (Figure 9). STILT can complement the adjoint approach by serving as a module for resolving the subgrid scale influence regions until dispersion has increased the influence area to a size which can be adequately resolved by grids in adjoint models [*Uliasz and Pielke*, 1990], where the information gain from adopting the particle method approaches 0.

[58] **Acknowledgments.** We would like to acknowledge Roland Draxler from the NOAA Air Resources Laboratory for providing the source code of the HYSPLIT model. The COBRA project was jointly funded by NSF (ATM-9821044), DOE (DE-FG02-98ER62695), NASA (NAG5-7950), and NOAA (NA06GP0406). J.C.L. was supported by the NASA Earth System Science Fellowship program.

## References

Andrews, A. E., K. A. Boering, B. C. Daube, S. C. Wofsy, E. J. Hintsa, E. M. Weinstock, and T. P. Bui, Empirical age spectra for the lower tropical stratosphere from in situ observations of $CO_2$: Implications for stratospheric transport, *J. Geophys. Res.*, *104*(D21), 26,581–26,595, 1999.

Bakwin, P. S., et al., Measurements of carbon dioxide on a very tall tower, *Tellus, Ser. B*, *47*, 535–549, 1996.

Belward, A. S., J. E. Estes, and K. D. Kline, The IGBP-DIS global 1-km land-cover data set DISCover: A project overview, *Photogramm. Eng. Remote Sens.*, *65*(9), 1013–1020, 1999.

Black, T. L., The new NMC mesoscale Eta-model—Description and forecast examples, *Weather Forecasting*, *9*, 265–278, 1994.

Boering, K. A., B. C. Daube Jr., S. C. Wofsy, M. Loewenstein, J. R. Podolske, and E. R. Keim, Tracer-tracer relationships and lower stratospheric dynamics: $CO_2$ and $N_2O$ correlations during SPADE, *Geophys. Res. Lett.*, *21*(23), 2567–2570, 1994.

Brillouin, L., *Science and Information Theory*, 320 pp., Academic, San Diego, Calif., 1956.

Byun, D. W., Dynamically consistent formulations in meteorological and air quality models for multiscale atmospheric studies, part II, Mass conservation issues, *J. Atmos. Sci.*, *56*, 3808–3820, 1999.

Chou, W. W., S. C. Wofsy, R. C. Harriss, J. C. Lin, C. Gerbig, and G. W. Sachse, Net fluxes of $CO_2$ in Amazonia derived from aircraft observations, *J. Geophys. Res.*, *107*(D22), 4614, doi:10.1029/2001JD001295, 2002.

Conway, T. J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang, Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res.*, *99*(D11), 22,831–22,855, 1994.

Cotton, W. R., et al., Cloud venting—A review and some new global annual estimates, *Earth Sci. Rev.*, *39*, 169–206, 1995.

Daube, B. C., et al., A high-precision fast-response airborne $CO_2$ analyzer for in situ sampling from the surface to the middle stratosphere, *J. Atmos. Oceanic Technol.*, *9*, 1532–1543, 2002.

Denmead, O. T., et al., Boundary layer budgets for regional estimates of scalar fluxes, *Global Change Biol.*, *2*, 255–264, 1996.

Denning, A. S., et al., Simulations of terrestrial carbon metabolism and atmospheric $CO_2$ in a general circulation model, part I, Surface carbon fluxes, *Tellus, Ser. B*, *48*, 521–542, 1996.

Draxler, R. R., and G. D. Hess, An overview of the HYSPLIT_4 modeling system for trajectories, dispersion, and deposition, *Aust. Meteorol. Mag.*, *47*, 295–308, 1998.

Errico, R. M., What is an adjoint model?, *Bull. Am. Meteorol. Soc.*, *78*(11), 2577–2591, 1997.

Fay, B., et al., Evaluation of Eulerian and Lagrangian atmospheric transport models at the Deutscher Wetterdienst using ANATEX surface tracer data, *Atmos. Environ.*, *29*, 2485–2497, 1995.

Giering, R., Tangent linear and adjoint biogeochemical models, in *Inverse Methods in Global Biogeochemical Cycles*, *Geophys. Monogr. Ser.*, vol. 114, edited by P. Kasibhatla et al., pp. 33–48, AGU, Washington, D. C., 1999.

Hanna, S. R., Some statistics of Lagrangian and Eulerian wind fluctuations, *J. Appl. Meteorol.*, *18*, 518–525, 1979.

Hanna, S. R., Applications in air pollution modeling, in *Atmospheric Turbulence and Air Pollution Modeling*, edited by F. T. M. Nieuwstadt and H. van Dop, 358 pp., D. Reidel, Norwell, Mass., 1982.

Heimann, M., and P. S. Kasibhatla, Introduction and overview, in *Inverse Methods in Global Biogeochemical Cycles*, *Geophys. Monogr. Ser.*, vol. 114, edited by P. Kasibhatla et al., pp. 1–2, AGU, Washington, D. C., 1999.

Hoerling, M. P., and L. L. Sanford, On the uncertainty in estimates of atmospheric heating due to data postprocessing, *J. Clim.*, *6*, 168–174, 1993.

Hoke, J. E., et al., The regional analysis and forecast system of the National Meteorological Center, *Weather Forecasting*, *4*, 323–334, 1989.

Holzer, M., and T. M. Hall, Transit-time and tracer-age distributions in geophysical flows, *J. Atmos. Sci.*, *57*, 3539–3558, 2000.

Kuck, L. R., et al., Measurements of landscape-scale fluxes of carbon dioxide in the Peruvian Amazon by vertical profiling through the atmospheric boundary layer, *J. Geophys. Res.*, *105*(D17), 22,137–22,146, 2000.

Lamb, R. G., H. Hogo, and L. E. Reid, A Lagrangian approach to modeling air pollutant dispersion: Development and testing in the vicinity of roadway, *Tech. Rep. 600/4/79-023*, Environ. Prot. Agency, Washington, D. C., 1979.

Legg, B. J., and M. R. Raupach, Markov-chain simulations of particle dispersion in inhomogeneous flows: The mean drift velocity induced by a gradient in Eulerian velocity variance, *Boundary Layer Meteorol.*, *24*, 3–13, 1982.

Levy, P. E., et al., Regional-scale $CO_2$ fluxes over central Sweden by a boundary layer budget method, *Agric. For. Meteorol.*, *98–99*, 169–180, 1999.

Lloyd, J., et al., Vertical profiles, boundary layer budgets, and regional flux estimates for $CO_2$ and its $^{13}C/^{12}C$ ratio and for water vapor above a forest/bog mosaic in central Siberia, *Global Biogeochem. Cycles*, *15*(2), 267–284, 2001.

Luhar, A.K., and R.E. Britter, A random walk model for dispersion in inhomogeneous turbulence in a convective boundary layer, *Atmos. Environ.*, *23*, 1911–1924, 1989.

Mooney, H. A., P. M. Vitousek, and P. A. Matson, Exchange of materials between terrestrial ecosystems and the atmosphere, *Science*, *238*, 926–932, 1987.

Obukhov, A. M., Description of turbulence in terms of Lagrangian variables, *Adv. Geophys.*, *6*, 113–116, 1959.

Odman, M. T., A quantitative analysis of numerical diffusion introduced by advection algorithms in air quality models, *Atmos. Environ.*, *31*, 1933–1940, 1997.

Pudykiewicz, J. A., Application of adjoint tracer transport equations for evaluating source parameters, *Atmos. Environ.*, *32*, 3039–3050, 1998.

Renyi, A., *A Diary on Information Theory*, 192 pp., Akad. Kiado, Budapest, 1984.

Rodean, H. C., *Stochastic Lagrangian Models of Turbulent Diffusion*, 84 pp., Am. Meteorol. Soc., Boston, Mass., 1996.

Ryall, D. B., et al., Estimating source regions of European emissions of trace gases from observations at Mace Head, *Atmos. Environ.*, *35*, 2507–2523, 2001.

Shannon, C., and W. Weaver, *The Mathematical Theory of Communication*, 125 pp., Univ. of Ill. Press, Urbana, 1963.

Smith, F. B., Conditioned particle motion in a homogeneous turbulent field, *Atmos. Environ.*, *2*, 491–508, 1968.

Stephens, B. B., et al., The $CO_2$ budget and rectification airborne study: Strategies for measuring rectifiers and regional fluxes, in *Inverse Methods in Global Biogeochemical Cycles*, *Geophys. Monogr. Ser.*, vol. 114, edited by P. Kasibhatla et al., pp. 311–324, AGU, Washington, D. C., 1999.

Stohl, A., Computation, accuracy and applications of trajectories—A review and bibliography, *Atmos. Environ.*, *32*, 947–966, 1998.

Stohl, A., and D. J. Thomson, A density correction for Lagrangian particle dispersion models, *Boundary Layer Meteorol.*, *90*, 155–167, 1999.

Stohl, A., and G. Wotawa, A method for computing single trajectories representing boundary layer transport, *Atmos. Environ.*, *29*, 3235–3238, 1993.

Stohl, A., G. Wotawa, P. Seibert, and H. Kromp-Kolb, Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories, *J. Appl. Meteorol.*, *34*, 2149–2165, 1995.

Stohl, A., M. Hittenberger, and G. Wotawa, Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data, *Atmos. Environ.*, *32*, 4245–4264, 1998.

Thomson, D. J., Random walk modelling of diffusion in inhomogeneous turbulence, *Q. J. R. Meteorol. Soc.*, *110*, 1107–1120, 1984.

Thomson, D. J., Criteria for the selection of stochastic models of particle trajectories in turbulent flows, *J. Fluid Mech.*, *180*, 529–556, 1987.

Thomson, D. J., W. L. Physick, and R. H. Maryon, Treatment of interfaces in random walk dispersion models, *J. Appl. Meteorol.*, *36*, 1284–1295, 1997.

Trenberth, K. E., Climate diagnostics from global analyses: Conservation of mass in ECMWF analyses, *J. Clim.*, *4*, 707–722, 1991.

Uliasz, M., Regional modeling of air pollution transport in the southwestern United States, in *Environmental Modeling: Computer Methods and Software for Simulating Environmental Pollution and its Adverse Effects*, edited by P. Zannetti, pp. 145–181, Comput. Mech., Billerica, Mass., 1996.

Uliasz, M., and R. A. Pielke, Receptor-oriented Lagrangian-Eulerian model of mesoscale air pollution dispersion, in *Computer Techniques in Environmental Studies*, edited by P. Zannetti, pp. 57–68, Comput. Mech., Billerica, Mass., 1990.

Vogelezang, D. H. P., and A. A. M. Holtslag, Evaluation and model impacts of alternative boundary-layer height formulations, *Boundary Layer Meteorol.*, *81*, 245–269, 1996.

Wilson, J. D., and Y. Zhuang, Restriction on the timestep to be used in stochastic Lagrangian models of turbulent dispersion, *Boundary Layer Meteorol.*, *49*, 309–316, 1989.

Yi, C., K. J. Davis, P. S. Bakwin, B. W. Berger, and L. C. Marr, Influence of advection on measurements of the net ecosystem-atmosphere exchange of $CO_2$ from a very tall tower, *J. Geophys. Res.*, *105*(D8), 9991–9999, 2000.

Zannetti, P., *Air Pollution Modeling*, 444 pp., Van Nostrand Reinhold, New York, 1990.

A. E. Andrews, NASA Goddard Space Flight Center, Mailstop 916.0, Greenbelt, MD 20771, USA. (andrews@code916.gsfc.nasa.gov)

B. C. Daube, C. Gerbig, J. C. Lin, and S. C. Wofsy, Department of Earth and Planetary Sciences, Harvard University, Pierce Hall, 29 Oxford Street, Cambridge, MA 02138, USA. (bcd@io.harvard.edu; chg@io.harvard.edu; johnlin@fas.harvard.edu; scw@io.harvard.edu)

K. J. Davis, Department of Meteorology, Pennsylvania State University, 512 Walker Building, University Park, PA 16802-5013, USA. (davis@essc.psu.edu)

C. A. Grainger, Department of Atmospheric Sciences, University of North Dakota, Box 9006, Grand Forks, ND 58202-9006, USA. (grainger@aero.und.edu)

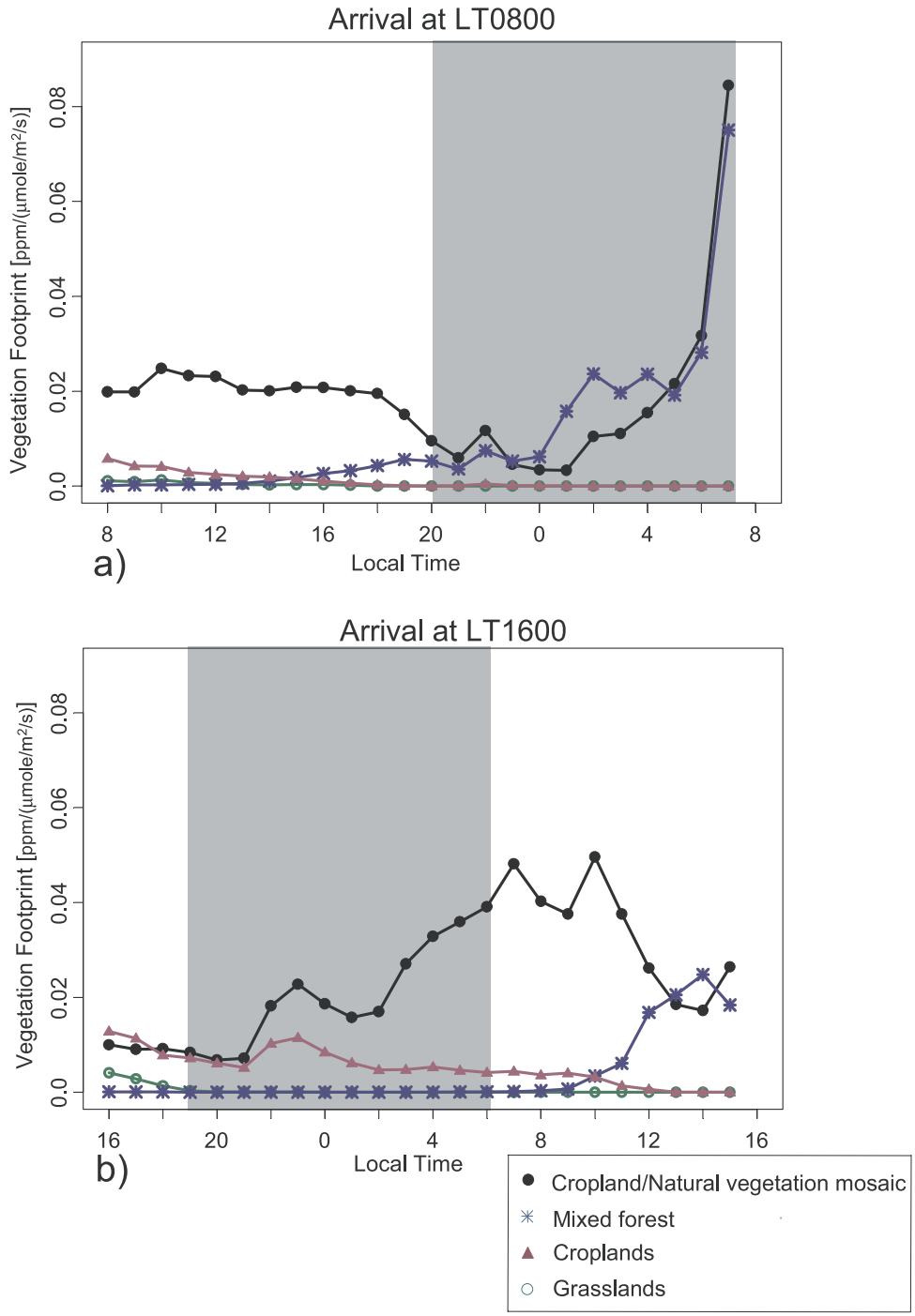


**Figure 8.** Time series of footprints due to the four main vegetation classes during the 24 hours prior to arrival on 8 June 1999 at (a) 0800 LT and (b) 1600 LT. Shaded portions denote nighttime periods. At each time step, particle locations were mapped onto the IGBP 1 km resolution vegetation data, regridded to 1/6° latitude by 1/4° longitude resolution. At the particle locations, footprints from different vegetation classes were quantified according to equation (8), with the time interval $\tau$ set to 1 hour. An unit flux of 1 $\mu$mol $m^{-2}$ $s^{-1}$ by a particular vegetation class persisting over an hour would lead to the change in $CO_2$ concentration at WLEF specified by the vegetation footprint.