ABSTRACT
Over the last 100 years, boundary layer meteorology grew from the subject of mostly near-surface observations to a field encompassing diverse atmospheric boundary layers (ABLs) around the world. From the start, researchers drew from an ever-expanding set of disciplines—thermodynamics, soil and plant studies, fluid dynamics and turbulence, cloud microphysics, and aerosol studies. Research expanded upward to include the entire ABL in response to the need to know how particles and trace gases dispersed, and later how to represent the ABL in numerical models of weather and climate (starting in the 1970s–80s); taking advantage of the opportunities afforded by the development of large-eddy simulations (1970s), direct numerical simulations (1990s), and a host of instruments to sample the boundary layer in situ and remotely from the surface, the air, and space. Near-surface flux-profile relationships were developed rapidly between the 1940s and 1970s, when rapid progress shifted to the fair-weather convective boundary layer (CBL), though tropical CBL studies date back to the 1940s. In the 1980s, ABL research began to include the interaction of the ABL with the surface and clouds, the first ABL parameterization schemes emerged; and land surface and ocean surface model development blossomed. Research in subsequent decades has focused on more complex ABLs, often identified by shortcomings or uncertainties in weather and climate models, including the stable boundary layer, the Arctic boundary layer, cloudy boundary layers, and ABLs over heterogeneous surfaces (including cities). The paper closes with a brief summary, some lessons learned, and a look to the future.

1. Introduction

How do we define the atmospheric boundary layer (ABL)? Here are some definitions from textbooks:

“that part of the troposphere that is directly influenced by the presence of the earth’s surface, and responds to surface forcings with a timescale of about an hour or less” (Stull 1988, p. 2).

“the layer of air directly above the earth’s surface in which the effects of the surface (friction, heating, and cooling) are felt directly on time scales less than a day, and in which significant fluxes of momentum, heat, or matter are carried by turbulent motions on a scale of the order of the depth of the boundary layer or less” (Garratt 1992, p. 1).

“lowest kilometer” or “lowest portion of the atmosphere, which intensively exchanges heat as well as mass and momentum with the earth’s surface” (Sorbjan 1989, p. 1).

The time scales in these definitions are fundamental, but vary by an order of magnitude, a function of context and application. On the short end, it takes minutes for air in a buoyant updraft to rise from a strongly heated surface to 1000 m above ground, where it can deposit pollutants or form a cloud; over a 24-h period, the temperature in the boundary layer varies more strongly than the air higher...
up, which can lead to killer frosts on clear nights with weak wind. Fields of wind, temperature, pressure, and other flow variables in the lowest atmosphere are mostly turbulent—complex, three-dimensional, irregular, and stochastic. Turbulent motions in the ABL drive the exchange of momentum, heat, and moisture between the surface and the atmosphere, and strongly affect human activities.

“Boundary layer” has a specific meaning in theoretical and experimental fluid dynamics, and the ABL is both similar to and different from those boundary layers. It is similar in being directly affected by a surface, but different in scale (on the order of a kilometer rather than a meter or less). Laboratory flows have well-defined boundary conditions, while Earth’s heterogeneous surface, variation in radiation, and weather provide conditions that are less well behaved. Nevertheless, it is convenient to divide the boundary layer into an inner layer (the “surface” layer) and outer layers, using the vertical stratification of temperature, wind, and turbulence intensity.

The textbooks provide a crude definition of the surface layer, which “occupies the lowest 10% of the whole ABL, which is about 10^9 to 10^12 m” (Sorbjan 1989); and is “the region at the bottom of the boundary layer where turbulent fluxes and stresses vary by less than 10% of their magnitude.” (Stull 1988, p. 10). The two definitions are equivalent for the commonly studied convective boundary layer (CBL), for which the fluxes vary almost linearly with height.

Here, we mostly restrict ourselves to the fair-weather ABL, following the tradition of most boundary layer research. We start with the “initial conditions,” recognizing the achievements of classical chemistry and physics, followed by those of early meteorologists often motivated by agriculture, and early fluid dynamicists who focused on turbulence and turbulent boundary layers, often in support of aircraft design, and follow with the story of the development of large-eddy simulation starting in the late 1960s. The discoveries and course corrections in our learning over the last 100 years were also enabled by development of instrumentation, both in the field and in the laboratory, and summarized in Stith et al. (2019). The material that follows is organized by topic, starting with the surface layer over land and ocean, and then moving upward to the entire ABL and its evolution under “ideal” (fair weather, no cloud) conditions. The complications of a heterogeneous surface and clouds follow, along with a description of the boundary layer over the Arctic Ocean, because we can learn from its unique characteristics. This is followed by a brief summary of how our knowledge of surface, boundary layer, and cloud processes is synthesized in terms of dispersion models, and the land surface and ABL parameterizations used in weather and climate models. We close with a historical synopsis, some lessons learned, and challenges and opportunities for the future.

This narrative is by its nature incomplete, reflecting the interests and backgrounds of its authors; and our identification of “first” discoveries should be taken with caution. We refer interested readers to a few books used heavily in preparation in this manuscript, in particular A Voyage through Turbulence, by Davidson et al. (2011). The Climate near the Ground, by Geiger (1966); and Meteorology: A Historical Survey, Vol. I, by Khrgian (1959, translated from Russian in 1970). We also consulted many textbooks along the way; most useful was Evaporation into the Atmosphere, by W. Brutsaert (1982), because of its emphasis on history. Finally, The Thermal Theory of Cyclones, by G. Kutzbach (1979) provided some surprising information about how much was known during the 1800s.

2. Historical origins

a. Early background

Our earliest knowledge of surface processes derives from agriculture. Farmers were certainly aware, for example, that killing frosts at night were more likely in low-lying areas. Indeed, McAdie (1915) discusses the association of cool air with lower elevations and the role of cold-air drainage, but more as a motivation for the real theme of his paper, entitled “The Theory and Practise (sic) of Frost Fighting.” At the same time, scientists were exploring the relationship of temperature with height at different times of the day. For example, Dines (1882) compared maximum and minimum temperatures at around 5 and 50 ft, consistently finding average 5-ft minimum temperatures about a degree Fahrenheit (°F) lower than at 50 ft, and average maximum temperatures about 0.5°–2.3°F higher, with the largest differences under clear skies. Based on observing nighttime temperature drops above and below the top of fog banks, he concluded that fog-bank tops, like grasses under clear skies, cool as a result of outgoing infrared radiation. The necessary measurements to verify this behavior were available not long after: the first longwave

---

1 Readers interested in exploring the differences further are referred to Nieuwstadt (1998).

2 We turn our attention to the Arctic rather than the poles because its boundary layer is over an ocean with ice as well as subject to high-latitude conditions, which makes it truly unique. Also, there appears to be far less written about the Antarctic boundary layer.
radiometers were invented in the 1890s (Hansch et al. 2006). Solar radiation was being measured by the mid-nineteenth century. Both measurements were likely used in agricultural research shortly after 1900 (Geiger 1966).

The concepts of the ideal gas law, the adiabatic process, and the release of latent heat by condensing water were all developed by 1841, when James Pollard Espy described how a plume of air, heated by Earth’s surface, rises to form a cloud in his *Philosophy of Storms* (Kutzbach 1979). Described in terms familiar today, he set the stage for describing how the daytime boundary layer heats up, by pointing out that a parcel of air, rendered buoyant by heating and evaporation at the surface, cools dry adiabatically as it rises, until it reaches the condensation level, where the cooling is slowed by the release of latent heat. The onset of manned balloon flights in the 1700s and 1800s made many aware that the temperature changes with height, with more systematic measurements after around 1850–70 (Kutzbach 1979).

Pomortsev used a series of manned-balloon training flights in Russia between 1885 and 1897 to document the change of wind with height until it was parallel to the isobars, and he associated that change with friction (proportional to the square of the wind speed). He also observed sudden changes in temperature and humidity (likely at the ABL capping inversion), and a dropoff of temperature with height at close to the adiabatic lapse rate beneath (Khrgian 1959, p. 270).

**b. Early development of turbulence theory**

The onset of contemporary boundary layer meteorology benefitted greatly from the developments in fluid dynamics around the beginning of the twentieth century. The term *boundary layer* was introduced by the German engineer Ludwig Prandtl (1905), who demonstrated that the flow around a solid body can be divided into two regions: a thin layer close to the body, where surface friction plays an essential role, and the external region, where the influence of the body’s surface can be neglected. On the other hand, the existence of the two distinctively different types of fluid motion—laminar and turbulent—was recognized by the first half of the nineteenth century. As for the term *turbulence*, it was likely first proposed in the fluid-mechanical context in 1887 by Sir William Thomson (Lord Kelvin; Thomson 1887), a prominent British physicist and mathematician. Joseph Boussinesq, a French mathematician and physicist, who introduced the concept of turbulent (eddy) diffusivity widely used in atmospheric boundary layer theory and modeling, described turbulent fluctuations of velocity as tumultuous motions or eddy agitations in his 1897 book (Boussinesq 1897). Foundations of modern turbulence theory were laid at the end of the nineteenth century by the British mathematician and engineer Osborne Reynolds. In his pioneering works of 1883 and 1895, Reynolds focused on conditions that determine laminar versus turbulent character of fluid flows in pipes and proposed a dimensionless parameter (Reynolds number Re, which is named after him) to be used as a criterion for dynamic similarity of viscous incompressible flows (Reynolds 1883, 1895). This number may be interpreted as the ratio of kinetic energy production by inertial forces to the energy dissipation by viscous forces and is often written as Re = UD/ν, where U is a characteristic flow speed, D a characteristic flow dimension (length scale), and ν the molecular kinematic viscosity. Flows with sufficiently high Re are expected to be turbulent, and flows with sufficiently low Re are laminar. In the spirit of statistical description of turbulent flows, Reynolds also suggested separating turbulent flow variables into the slowly varying mean and highly variable fluctuating parts (components) in order to study turbulent flow in terms of mean-flow variables and statistics of turbulent fluctuations of various orders. Reynolds’s original approach to evaluating turbulence statistics was based on averaging flow fields over a spatial length (or volume), according to certain rules. These rules were formulated in a way that later facilitated their implementation in the Reynolds-averaged equations of fluid dynamics extensively used in boundary layer meteorology. The presently adopted concept of turbulence averaging, which is understood as ensemble averaging, differs considerably from the averaging approach put forward by Reynolds, but remarkably preserves the Reynolds averaging properties.

The story of the publication of Reynolds’s seminal 1895 paper was not as glorious as one might expect. Sir George Stokes and Dr. Horace Lamb (he was just a professor at that time), both very prominent figures in fluid dynamics in those days, were approached by the *Transactions of the Royal Society* editor Lord Rayleigh to review the article. They were rather lukewarm in their “unenthusiastic, even grudging, acquiescence to the paper’s acceptance for publication” (Lauder 2015). To a great extent, this lack of appreciation is thought to

---

3 According to Ansell et al. (2006), several sites were taking pressure measurements in the late 1800s, providing some plausibility to at least having surface pressure analyses available.

4 Leonardo de Vinci anticipated the existence of turbulence 400 years earlier, describing flow as a “principal current” and “random and reverse motions” (Richter 1970, Plate 20 and Note 389).

5 Oberbeck (1888) also qualitatively describes an eddy viscosity, mentioning “currents of a local nature which will increase the (internal) friction” (p. 164 of English translation), suggesting the idea was around.
be due to Reynolds’s excessively wordy style of writing, his overfocus on well-established procedures and principles while leaving his novel ideas unexplained, and his tendency to use a self-invented terminology that was puzzling to the referees. The long-term effect of the paper was nevertheless tremendous, as it outlined a practical and significant approach to study turbulent flows both experimentally and theoretically.

Interestingly, during the same year Reynolds’s paper on turbulence decomposition and averaging went to print, Lamb published his Hydrodynamics (1895), which is regarded as the first fluid-mechanics textbook. The notion of turbulent flow appears toward the very end of the book, with the author presenting a review of Reynolds’s experiments in pipes and pointing to principal differences between laminar and turbulent flows, calling the latter “wildly irregular” with “interlacing and constantly varying streams, crossing and recrossing the pipe.” Shortly before his death in 1934, Sir Horace Lamb wittily expressed the difficulty of explaining and studying turbulence by reportedly saying in his address to the British Association for the Advancement of Science: “I am an old man now, and when I die and go to Heaven, there are two matters on which I hope to be enlightened. One is quantum electrodynamics, and the other is turbulence. About the former, I am really rather optimistic” (Goldstein 1969).

The first analytical treatment of boundary layer flow in the geophysical context is attributed to Vagn Ekman (1905), a Swedish oceanographer, who employed momentum balance equations for viscous fluid to explain Fritjof Nansen’s observation during the Fram expedition (Mohn 1905) that the surface current carried sea ice in a direction 20°–40° to the right of the prevailing wind direction. Although the problem setup was not particularly meteorological, Ekman mentioned possible atmospheric applications of his model, pointing to the analogy between the three-way balance among the pressure-gradient, Coriolis, and friction forces in the atmospheric and oceanic boundary layers. Ekman’s treatment of turbulence (specifically, turbulent friction) in his equations was rather straightforward. Based on the consideration that transfer of momentum between moving layers of fluid “owing to the irregular formation of eddies” is more “intense” than “due to friction alone,” he introduced a virtual viscosity with a value “much greater” than the actual molecular viscosity. (It is quite possible that Ekman was not aware of Boussinesq’s introduction of eddy viscosity.) Ekman also employed the momentum balance equations in their original form, that is without applying any decomposition of flow fields in mean and fluctuating components, and without subsequent averaging of equations following Reynolds or in some other manner. An approach very similar to that of Ekman was later used to describe atmospheric boundary layer wind profiles in the work of Filip Åkerblom (1908), another Swedish geophysicist, who never mentioned Ekman’s model in his paper but cited Lamb (1895) while deriving the governing equations. Carl-Gustaf Rossby, however, states in his 1927 review article (Rossby 1927) that Åkerblom was led to his study by Ekman’s (1905) paper.

The view of turbulent exchange in the atmosphere as a “process like molecular conduction but much more vigorous” was further advanced by Geoffrey Taylor, a British mathematician and physicist, in his 1915 paper about eddy motion in the atmosphere (Taylor 1915). He put forward the idea regarding “coherent fluid masses” that “move a certain distance up or down vertically carrying all their transferable properties and then mix with the surroundings in which they find themselves.” The name for such a distance, the “mixture length” (it transformed with time into “mixing length” in English scientific literature) was introduced in 1925 by Ludwig Prandtl, who allegedly was not aware of Taylor’s work (Wyngaard 2010). Interestingly, Taylor abandoned his own mixing-length concept in his famous 1921 paper on the nature of turbulent diffusion, where he pointed to the crucial role of covariances of turbulent fluctuations of different field variables as measures of transport of substances in a turbulent flow. In Prandtl’s works, however, the notion of mixing length was exploited extensively and was used to lay the foundation for a class of approaches that we currently call semiempirical turbulence theories (Monin and Yaglom 1971). The advancement of these theories, besides Prandtl and Taylor, was strongly influenced by Prandtl’s student Theodore von Kármán.

The semiempirical theories, which form the basis for the logarithmic profile developed by von Kármán (1931) to describe the flow of neutrally stratified air as a function of height and later the flux-profile relationships discussed in section 3, are based on the analogy between turbulence and random molecular motion. Besides using notions of mixing length (analogous to the molecular free path) as the turbulence length scale, and turbulent diffusivities (analogous to molecular diffusivities), semiempirical relations can be constructed between the fluxes of momentum and scalars and their corresponding gradients. These relationships have been derived from observations of turbulent flow in the laboratory and in the atmosphere, complemented with considerations of dimensionality and symmetry, plausible physical arguments, scale analysis, and similarity theories.

Semiempirical theories also played an especially important role in seeking practical solutions to the so-called turbulence closure problem as applied to the
ABL parameterization schemes discussed in section 10. This problem finds its origin in the chain of differential equations for statistical moments of turbulent flow fields derived by Keller and Friedmann (1924), referred in the literature as Friedmann–Keller equations, in which expressions for the low-order turbulence moments unavoidably contain terms with higher-order moments. Thus, the sequence of the equations appears to be infinite, as any individual subset of these equations remains unclosed, that is, it contains more unknowns than equations available to solve for these unknowns. Semiempirical methods and hypotheses were essential in formulation of physical assumptions that enabled closing the lowest-order equations from the Friedmann–Keller set containing only first- (mean fields) and second-order (variances and covariances) turbulence statistics in terms of eddy-exchange coefficients [the so-called Reynolds-Averaged Navier–Stokes (RANS) equations].

One more group of efforts dating back to the 1920s and 1930s focused on understanding the physical mechanism of turbulence. Lewis Richardson, a British mathematician and meteorologist, hypothesized in 1922 that turbulent flows are composed of eddies of different sizes, and that kinetic energy of the flow cascades from large to small scales of motion. According to Richardson, eddies of a particular size emerge because of instability of eddies of a larger size, acquire energy from these unstable eddies, lose stability themselves, and fall apart transferring energy to the smaller eddies. The largest-scale eddies are produced as a result of the instability of the mean motion. This cascading generation of smaller and smaller eddies continues until the eddies become small enough to feel the direct effect of viscosity, which plays a stabilizing role by preventing the smallest eddies from further destruction, while the transferred energy dissipates into heat. That was a fundamental insight in the mechanism of energy transfer in the turbulent flow, although Richardson did not come up with any mathematical formulation of his theory. In 1926, Richardson made a breakthrough contribution to the theory of turbulent diffusion by establishing through physical arguments that the effective diffusion coefficient for a plume of passive constituent in developed turbulence is proportional to the four-thirds power of the characteristic length scale of the plume (Richardson 1926). Until the seminal works of Kolmogorov and Obukhov in the early 1940s (discussed below), this was the only quantitative prediction regarding the physical behavior of small-scale turbulence (Monin and Yaglom 1971). Richardson (1920) also developed a criterion for turbulence breakdown in sheared, stratified flows known today as the Richardson number.

Another important contribution to conceptual understanding of small-scale turbulent motions was made in 1935 by Taylor, who introduced the notion of homogeneous and isotropic turbulence, for which the probability distributions of the fluid variables are the same, regardless of translation, reflection, or orientation of the coordinate system (Taylor 1935). Under assumptions of homogeneity and isotropy, turbulent flow becomes more mathematically tractable in terms of the Friedmann–Keller equations, despite the fact that all fundamental difficulties associated with structural properties of these equations remain intact. Although the assumptions of homogeneous and isotropic turbulence are never generally satisfied for any actual atmospheric turbulent flow, the mathematical tools developed for such an idealized turbulence state turned out to be useful to describe statistical turbulence structure in real flows (including atmospheric ones) at large Reynolds numbers, where small-scale turbulence may be considered locally homogeneous and locally isotropic.

The next big step in the development of turbulence theory in the first half of the twentieth century occurred when two Russian scientists—Andrey Kolmogorov (1941a,b) and Alexander Obukhov (1941)—laid the foundation for modern quantitative theory of the energy transfer processes in turbulent flows, which had paramount implications for fundamental and applied aspects of boundary layer meteorology. Kolmogorov significantly refined Richardson’s concept of the turbulence energy cascade in physical terms and put it on a solid mathematical basis. He pointed out that, no matter how inhomogeneous and anisotropic the large-scale structural flow elements are, the small-scale turbulent fluctuations in a developed turbulent flow with a very large Reynolds number would be homogeneous and isotropic. The fluctuations would also be quasi-stationary on the time scale of evolution of the mean flow. Based on such considerations, Kolmogorov formulated two hypotheses. First, that statistical properties of small-scale turbulence fluctuations are controlled by only two parameters: the kinematic viscosity of the fluid and the turbulence kinetic energy dissipation rate (which is also a measure of energy transport across the turbulence spectrum, i.e., from large to small turbulent eddies). Combinations of these two parameters provide length, velocity, and time scales of the smallest turbulent fluctuations that exist in a given turbulent flow (the so-called Kolmogorov microscales). With his second hypothesis, Kolmogorov postulated and physically substantiated the existence of the so-called inertial range of turbulence scales that are much larger than the Kolmogorov length scale, but much smaller than the scale of largest turbulent motions originating from instabilities of the mean flow. Within this range, the statistical regime
of turbulence is determined solely by the energy dissipation rate. For turbulence in the inertial range (or as we call it today, inertial subrange), Kolmogorov established the now famous two-thirds law for the scale dependence of the structure function of velocity fluctuations, and Obukhov derived the equivalent and even more widely known minus-five-thirds law for decay of velocity spectral density with wavenumber. This so-called $-5/3$ spectrum has since been applied in turbulence modeling, is often used to evaluate both measurements and large-eddy simulations for convective boundary layers, and provides a tool to estimate viscous dissipation.

c. Large-eddy simulation

Although the birth of large-eddy simulation (LES) occurred well into the twentieth century, its significance to boundary layer meteorology merits a brief treatment here. The eddy size in the ABL varies from a few kilometers down to a few millimeters—too much to resolve directly by even the most advanced computers today, so LES is designed to resolve a subset—the large eddies (of order of ABL depth down to tens of meters to meters)—that account for most of the turbulence kinetic energy and fluxes, while parameterizing the effects of the smaller eddies. By the middle of the twentieth century, development of electronic computers enabled the development of numerical weather prediction (NWP) and general circulation models (GCMs), (see Randall et al. 2019; Benjamin et al. 2019), providing a foundation for LES development.

The ready availability of such a powerful (for the time) computer enabled James W. Deardorff of the National Center for Atmospheric Research (NCAR) to develop the first LES in the late 1960s, with significant input from Douglas K. Lilly, who arrived at NCAR in 1964 with numerical-modeling insights he and colleagues developed at the General Circulation Laboratory (a predecessor of the Geophysical Fluid Dynamics Laboratory at Princeton; Kanak 2004). As in the case of NWP models and GCMs, the idea was to solve the rate equations for the relevant flow variables on a grid, while representing subgrid effects (smaller, unresolved scales) with some sort of parameterization. However, unlike NWP models and GCMs, the grid spacing used by LES was sufficiently fine that the simulations generated 3D time dependent random fields, that is, the LES solutions were stochastic in space and time. Early LES employed horizontal periodic boundary conditions (consistent with the assumption of horizontal homogeneity), effectively generating turbulent inflow and outflow conditions required by the computations.

Not long after, more modifications brought LES closer to the real atmosphere, with radiation included in Deardorff (1974a), and clouds along with the effects of horizontal advection and subsidence in Sommeria (1976). The domains for these early LES ranged from 2 to 5 km in the horizontal directions, and up to 2 km in the vertical direction, with horizontal grid spacing between 50 and 125 m, and the vertical spacing 50 m. Because of the finer resolutions needed to resolve the large eddies and limited computer capability, it took more than an additional decade until the first successful LES of the stable boundary layer was conducted by Mason and Derbyshire (1990).

One of the most challenging aspects of LES is accounting for the impacts of the subgrid-scale (SGS) eddies, those too small to be resolved. The first, and still most commonly used parameterizations assume that the SGS fluxes are proportional to the eddy-exchange coefficients for momentum $K_M$ and heat $K_H$. In his early LES work, Deardorff (1972a) used an expression for $K_M$ inspired by that derived by Smagorinsky (1963) for NWP/GCMs, with theoretical roots in the seminal work by Lilly (1966a). In the Smagorinsky–Lilly formulation, $K_M$ is a function of the strain rate tensor, the grid spacing, and a constant (the “Smagorinsky constant’’), with $K_H = 3K_M$, or equivalently, the turbulent Prandtl number $Pr_T = K_M/K_H = 1/3$. At the lowest grid level, however, Deardorff set $Pr_T \approx 1$ based on measurements described in Businger (1966), an early precedent to the universal practice of requiring the near-surface flow fields in LES to match Monin–Obukhov similarity theory (see section 3b). Lilly (1962) included the effects of buoyancy by allowing $K_M$ to vary with Richardson number in a simulation of a two-dimensional thermal. Deardorff (1980a) was the first to apply a turbulent kinetic energy (TKE) scheme for which $K_M$ is a function of the subgrid TKE [see (9-13), section 4], thus accounting for the effects of pressure fluctuations and turbulent transport as well as buoyancy, shear, and dissipation. In this SGS model, $Pr_T$ was weakly dependent on thermal stratification. Deardorff (1973) provides an insightful review of these early SGS model developments.

Most LES SGS models require that the grid size lies within the inertial subrange, so that turbulence could be treated as isotropic, something that applies best away from the lower boundary, very stable layers, or with sufficiently small grid spacing (e.g., Sullivan et al. 1994 and references therein). A more fundamental problem is

---

6 Which was anticipated by Richardson (1922).

7 Called the Deardorff TKE scheme in the review of LES by Moeng and Sullivan (2015), which credits Deardorff with this approach, likely inspired in part by discussions with Lilly, and earlier papers by Lilly (1966a, 1967).
the assumption that these representations are deterministic at each point, when in fact they are deterministic only in an average (or ensemble) sense. At any given time, energy at a grid point can be transported upscale (backscattered) rather than only downscale as implied in both the Smagorinsky–Lilly and Deardorff (1980a) TKE schemes by analogy to molecular transport. Mason and Thomson (1992) addressed this problem by modifying the Smagorinsky-type SGS scheme in their LES to include stochastic upscale flow of energy (backscatter) from the subgrid to resolved scales following the ideas of Leith (1990). An alternative, more phenomenological approach was proposed by Kosović (1997).

The origin of LES has roots in the atmospheric science community, but the engineering community rapidly adopted the technique (e.g., Reynolds 1976) and started to benefit the atmospheric boundary layer community, especially in the context of representing subgrid effects. Leonard (1974) and Germano (1986), for example, provided a useful way to represent SGS parameterization for the filtered Navier–Stokes equations, by dividing the subgrid stresses into three sets of terms that are Galilean invariant and independent of the filter used to separate resolved and subgrid scales. Findikakis and Street (1979) truncated the subgrid-scale rate equations to develop a set of algebraic equations for subgrid-scale heat and momentum transport as an alternative to the eddy-diffusion models. Deardorff (1980a) refers to the Findikakis–Street paper in interpreting his LES of a stratocumulus-topped boundary layer. Likewise, Moeng (1984), in her landmark paper that develops an LES solving the horizontal equations in spectral space, notes that solving the horizontal equations in this way facilitates handling at least partially the so-called Leonard (1974) stresses in the treatment of unresolved motions. The cross-field fertilization between the geophysical and engineering communities has continued to this day. In the United States, for example, many major papers on LES have come from engineering departments at Stanford, the University of Minnesota, and Johns Hopkins, among others, and several individuals with engineering backgrounds are working on LES at atmospheric science institutions or university departments.

Moeng (1984) believed that because “the large eddies are much more important and flow dependent than the small eddies,” that LES would be “relatively insensitive to the parameterization for the small eddies,” which was verified for the CBL based on a comparison of four different LES codes (Nieuwstadt et al. 1992). LES intercomparisons on modest meshes ~100^3 show larger differences for a neutral Ekman layer (Andrén et al. 1994) and a stable boundary layer (Beare et al. 2006); most of the differences were related to SGS parameterizations. This is of course largely because the turbulent eddies are smaller in the latter cases, so that the SGS representations are more exposed, especially near the lower boundary, where smaller eddies are more important. Sullivan and Patton (2011), using meshes of 1024^3, find LES solutions numerically converge provided there is adequate separation between the energy containing eddies and those near the filter cutoff scale.

An alternative type of SGS scheme involves the so-called dynamic approach, in which the LES fields are filtered using two (or more) different filters and variations between the filtered fields are used to derive the subgrid fluxes. Meneveau et al. (1996), Porté-Agel et al. (2000), and Bou-Zeid et al. (2005) developed a series of dynamic models using different averaging procedures in time or space to estimate the local Smagorinsky constant from the resolved-scale fields during the LES run rather than having to rely on a globally prescribed value, and Porté-Agel (2004) proposed a similar treatment for $K_f$. Eddy-diffusion schemes, however, must produce forward scattering. This shortcoming is avoided in the dynamic reconstruction method of Chow et al. (2005), through use of the resolved subfilter scales, and by Lu and Porté-Agel (2010), who develop a phenomenological dynamic approach. [For a brief discussion of some dynamic and other SGS schemes, see Shi et al. (2018b).] Ludwig et al. (2009) compared several SGS schemes for a neutral boundary layer, including some of the dynamic schemes, and found differences that show up even for the largest eddies, with the largest differences between the simple eddy-visibility schemes and dynamic models that include backscatter, and the dynamic schemes working the best. Similarly, Shi et al. (2018a) found dynamic schemes performed better in a simulation of a stratocumulus-topped boundary layer.

Poor resolution near the surface and its potential influence of large eddies higher up spurred the LES community to increase its focus on rigorous testing of SGS schemes (e.g., Meneveau and Katz 2000 and references therein). From the atmospheric perspective, field testing emerged in the late 1990s, largely spurred on by the work of a group at Pennsylvania State University (Tong et al. 1998, 1999), who showed how ABL

---

8 And perhaps the name “large-eddy simulation” as well. It was suggested by Parvis Moin to Wyngaard (2010, footnote on p. 136) that William Reynolds of Stanford coined the term.

9 Calculations do, however, sometimes produce negative eddy viscosities. When this occurs, the values are typically “clipped” (set to zero or a small negative value).
turbulence-scale interaction at small scales could be studied through sampling both filter/resolved and subfilter/subgrid scales of atmospheric variables using arrays of densely spaced instrumentation a few meters above the surface. The daunting task of setting up and maintaining the optimum arrays was beyond the reach of the early researchers, but efforts by the Pennsylvania State University group (Tong et al. 1999) and by a group from University of Minnesota and Johns Hopkins (Porté-Agel et al. 2001) were successfully sufficient to establish both the feasibility of this pioneering approach and its promise in relating filter-scale fields to subfilter fluxes (see also Higgins et al. 2003, 2004).

It was the technical staff at NCAR’s Atmospheric Technology Division who enabled researchers to fully implement the design set out by Tong et al. (1998) in the first of several Horizontal Array Turbulence Study [HATS (2000); Horst et al. 2004] field campaigns, building on the contributions of Tong et al. (1999) and Porté-Agel et al. (2001). Using HATS data, Sullivan et al. (2003) found that the Smagorinsky–Lilly and Deardorff TKE schemes were adequate if the turbulence length scale corresponding to the peak in the vertical velocity spectrum was large compared to the filter scale, but increasingly sophisticated methods using turbulence-spectrum treatment, can have important impact (e.g., Stevens et al. 1999).

The SGS problem is avoided in DNS, for which the time and space scales of all motions are resolved, making an SGS parameterization unnecessary. This of course requires much smaller Reynolds numbers and domains. For the atmosphere, with $L \sim 1000 \text{ m}$ and $U \sim 1–10 \text{ m s}^{-1}$, $Re = UL/\nu \sim 10^2–10^3$, since $\nu \sim 10^{-5} \text{ m s}^{-2}$. As in the case of LES, the first DNS was run on a computer at NCAR, by Steven Orszag of the Massachusetts Institute of Technology and Stuart Patterson of Swarthmore College (Orszag and Patterson 1972; also see review by Fox and Lilly 1972). For this first set of DNS runs, with $Re \approx 45$, the $32^3$-grid domain was designed to simulate wind-tunnel turbulence.

With today’s more powerful massively parallel computers, DNS using meshes of $2048^3$ and larger, extending to $Re \sim 10^4–10^7$, has become a popular tool to study idealized atmospheric boundary flows such as turbulent Ekman layers (Shah and Bou-Zeid 2014), entrainment into dry (García and Mellado 2014) or cloud-topped boundary layers (Mellado 2017), entrainment driven by shear (Jonker et al. 2013), the nocturnal jet (Fedorovich et al. 2017), katabatic flows (Shapiro and Fedorovich 2008), and turbulence “collapse” in the stable boundary layer (Ansorge and Mellado 2014). A review of the use of DNS as a research tool for turbulence can be found in Moin and Mahesh (1998).

As LESs became more sophisticated, they were increasingly used to improve the PBL parameterization...
schemes employed in weather and climate models (see section 10). As an example, early tests of what later became the Mellor–Yamada–Janjić (MYJ) PBL parameterization scheme, Yamada and Mellor (1975) used the Deardorff (1974a,b) simulations of the Wangara experiment (Clarke et al. 1971); and when Janjić (2001) described the MYJ scheme for the modeling community, he referred to several LES papers as sources of comparison or parameterization constants. LES played an even more central role in Nakanishi and Niino (2009), which eventually led to the Mellor–Yamada–Nakanishi–Niino (MYNN) PBL scheme. In contrast to these models, which are based on the turbulent variance- and flux-tendency equations, the Yonsei University (YSU) scheme (Hong et al. 2006) also draws heavily on results of LES as well as observations for, among other things, profiles of $K_H$ and $K_S$ (Noh et al. 2003). LES has also played an important role in developing the $K$-profile parameterization used in ocean models (Large et al. 1994). We will see in section 8 how the interplay of targeted observations and LES has led to significant progress in understanding and representing the stratocumulus-topped PBL, especially in the last two decades. Similarly, section 10a reveals the significant role of LES in dispersion studies.

Development of LES versions of mesoscale or numerical prediction models has provided additional new options for studying the evolving boundary layer over a heterogeneous surface, with a suite of physical parameterization options for clouds, radiation, the land surface, and the surface layer. Currently, many fine-grid weather or mesoscale models offer an LES option, perhaps the first of which was the Colorado State University (CSU) Regional Atmospheric Modeling System (RAMS) model (Pielke et al. 1992). Others have emerged since then, among which are the French Mes-oNH (Cuxart et al. 2000), the British Met Office Unified model (Beare and MacVean 2004), COAMPS (Hodur 1997), the Advanced Regional Prediction System (ARPS; e.g., Xue et al. 2001), COSMO (COSMO-LES; Loewe et al. 2017), and the Weather Research and Forecasting (WRF) Model (Mirocha et al. 2010).

The limitations associated with periodic boundary conditions remain an issue, but can be addressed by selecting a domain large enough not to constrain the scales of interest (e.g., de Roode et al. 2004) or by using so-called fringe methods to generate turbulent inflow and outflow conditions (Spalart and Watmuff 1993). Not surprisingly, as computer capability has improved, large-domain LESs have emerged [e.g., 205 km × 205 km in Khairoutdinov et al. (2009) to enable study of maritime deep convection and its environment], but such studies could suffer from effect of still larger scales. A third approach is to nest a finer-grid LES within larger, coarser LES domains (e.g., Mirocha et al. 2013). The next logical step—nesting down from mesoscale meshes to LES to study boundary layers influenced by real weather—is still in its infancy (see, e.g., Moeng et al. 2007; Zhu et al. 2010; Mirocha et al. 2014; Schalkwijk et al. 2015; Muñoz-Esparza et al. 2017).

### 3. The surface layer

As noted in the introduction, the surface layer is defined for convenience as having fluxes roughly independent of height, consistent with depths up to ~100 m in windy, convective conditions, and much shallower under stable conditions. For linear variation of fluxes with height, this amounts to about a 10% difference over the 10% of the BL depth often taken as the surface-layer depth. In the presence of surface roughness elements (stones, vegetation, buildings, waves, etc.) the surface layer lies above a roughness sublayer, where the effects of individual roughness elements are felt (e.g., Katul et al. 1999; Finnigan 2000), which translates to a depth roughly 2–5 times the height of the roughness elements (Raupach et al. 1991).

#### a. The surface energy budget

Surface-layer and PBL stability as well as the role of water vapor in buoyancy and cloud formation are affected by the near-surface temperature $T$ and specific humidity $q$ and their turbulent fluxes; these are in turn influenced by the surface energy budget (SEB). Thus representing the ABL in numerical models of weather and climate (section 10) requires accurate representation of the SEB. Neglecting effects of vegetation and advection, the SEB is given by

$$Q_{R0} = Q_{H0} + Q_{E0} + Q_{G0},$$

where the four terms are, respectively, the incoming net radiation, the sensible and latent heat flux from the surface to the air, and the heat flux into the ground (or surface), expressed in watts per meter squared. In practice, the terms are measured away from the surface and corrected to surface values assuming changes of heat and moisture content in the intervening air and soil layers are due to vertical flux divergence; with horizontal transport and vegetation effects often estimated.

By the 1920s, atmospheric radiation had been extensively studied, and people were estimating the albedo of different surfaces, and the effects of sun angle and atmospheric water vapor content were documented in the 1930s, according to Geiger (1966). Diurnal soil temperature variation was documented before 1900, as illustrated by Fig. 9-1, and several groups were sampling soil temperatures through the year. Soil moisture was likely...
estimated by drying soil samples, although de Vries (1953) attempted to estimate soil moisture from soil thermal conductivity, which was measured at Potsdam, Germany, by Albrecht (1937); and it was already recognized that both soil thermal conductivity and specific heat were functions of soil composition and water content.

The 1903 measurements of $Q_{R0}$ and $Q_{G0}$ at Potsdam enabled Albrecht (1940) to estimate $Q_{H0}$ and $Q_{E0}$ on a daily time scale from (9-1), quite possibly using Bowen's (1926) paper relating their ratio to vertical gradients of temperature and water vapor content. By the time of the 1953 Great Plains Turbulence field program (Lettau and Davidson 1957), better soil flux, moisture, temperature, and thermal-conductivity measurements were available and used to make estimates of soil heat flux and correct it to the surface soil heat flux $Q_{G0}$. Similarly, numerous radiometers sampled $Q_{R0}$. With these data available, Suomi used (9-1) and applied Bowen's assumptions to the vertical $\bar{v}$ and $\bar{q}$ gradients to obtain $Q_{H0}$ and $Q_{E0}$ (Lettau 1957).

According to Lettau (1957), Halstead, on the other hand, obtained the vertical fluxes of $\theta$ and $q$ from the bulk aerodynamic technique (e.g., Taylor 1916), which uses measurements at a single level (usually 10 m) and estimates (or measurements) of surface variables, such that turbulent fluxes $(\bar{w}^\theta)_0$, $(\bar{w}^q)_0$, and $-\bar{u}\bar{w} = u^*_w$, where $q$ is the specific humidity, $\theta$ is temperature $T$ corrected for height using the adiabatic lapse rate,$^{11}$ $u^*_w$ is called the friction velocity, and $\overline{U}$ is the mean wind speed, are given by

$$u^*_w = C_D \overline{U}^2,$$  \hspace{1cm} (9-2a)

$$\overline{(wT)}_0 = \overline{(w\theta)}_0 = C_H (\bar{q}_0 - \theta) \overline{U}, \hspace{0.5cm} \text{and} \hspace{0.5cm} \overline{(wq)}_0 = C_E (\bar{q}_0 - \bar{q}) \overline{U},$$  \hspace{1cm} (9-2b)

where the primes denote departures from time or horizontal averages (overbar), $C_D$ is the drag coefficient, $C_H$ and $C_E$ are the exchange coefficients for heat and moisture, and the subscript 0 again indicates a “surface value” (actually a small distance above the surface, see section 3b). Halstead assumed a logarithmic $\overline{U}$ profile, and obtained $C_H$ and $C_E$ from the molecular Prandtl number ($Pr = C_m/C_H = 0.71$), the molecular Schmidt number ($Sc = C_m/C_E = 1.64$), and $C_D$, which was based on drag-plate stress measurements. Unfortunately, (9-1) did not balance for Halstead.

By the 1960s, hot-wire anemometers and fine-wire resistance thermometers (e.g., Swinbank 1951); fast-response Lyman-alpha sensors, dewpoint hygrometers, and wet-bulb thermometers (Miyake and McBean 1970; Pond et al. 1971); and the development of the sonic anemometer (Schotland 1959; Barrett and Suomi 1949) made possible calculating the near-surface terms $\overline{wT}$ and $\overline{wj}$ directly (eddy correlation) to obtain $Q_{E0} = \bar{p}_e L_e (\overline{wj})_0$ and $Q_{H0} = \bar{p}_e C_p (\overline{wT})_0$, where $\rho$ is the air density, $L_e$ is the latent heat of vaporization, and $C_p$ is the specific heat at constant pressure; and measurements required for estimates of $Q_{G0}$ and $Q_{R0}$ continued to improve. But the observed surface energy budget still failed to balance. This prompted a workshop (Foken and Oncley 1995) to examine sources of the discrepancy, including impact of sonic-anemometer flow distortion and undersampling of eddy-correlation fluxes, the sampling of different areas for the estimates of the four terms in (9-1), inherent error in the instruments involved, heat storage between the surface and flux-measurement height, storage in the soil layer, and horizontal advection. The workshop resulted in the Energy Balance Experiment (EBEX) in a periodically flooded cotton field, in which the researchers tried to minimize the known sources of error; but there was still a daytime residual of about 10% (Fig. 9-2).

How can we eliminate that last 10%? Heusinkveld et al. (2004), van de Wiel et al. (2003), Liebenthal et al. (2005), and Steeneveld et al. (2006) focused on soil measurements to improve $Q_{G0}$, with some success in improving balance. However, others believe that observations of $Q_{H0}$ and $Q_{R0}$ are the culprits, suggesting corrections by multiplying them by the factor needed to force balance (e.g., Twine et al. 2000), sampling large eddies better,
either through long averaging time (Finnigan et al. 2003) or using distributed arrays (e.g., Engelmann and Bernhofer 2016), distributing flux sites to sample different types of land cover in heterogeneous areas (e.g., Kustas et al. 2005),12 mapping surface fluxes from aircraft (Mauder et al. 2008),12 or performing more rigorous corrections to correct for “shadowing effects” on sonic anemometers (Horst et al. 2015; Frank et al. 2016). Though some of the successes might be site or situation specific, a common theme in attempts to achieve balance is using the best instruments possible, installing them carefully, and doing thorough corrections afterward; building on the careful procedures used in EBEX.

b. Flux-profile relationships

In addition to the bulk aerodynamic technique, fluxes in the surface layer are estimated using the flux-profile technique:

\[ u^* = K_w \frac{\partial U}{\partial z}, \]  

(9-3a)

\[ w'T' = -K_w \frac{\partial T}{\partial z}, \]  

(9-3b)

\[ w'q' = -K_q \frac{\partial q}{\partial z}, \]  

(9-3c)

Starting in the mid-1900s, these relationships provided a theoretical basis for estimating the exchange coefficients in (9-2), drawing from the work discussed in section 2. What we now call Monin–Obukhov similarity theory (MOST) was developed during and just after World War II (Obukhov 1946, 1971; Monin and Obukhov 1954).13 Obukhov’s (1946) paper introduced the length scale \( L \), given by

\[ L = \frac{u^3 T}{kg w T}, \]  

(9-4)

where \( k \) is the von Kármán constant and \( g \) the acceleration of gravity. The 1954 paper uses dimensional arguments14 to demonstrate that temperature and wind speed profiles are universal functions of \( z/L \). Today, virtual temperature \( T_v \) is used instead of \( T \). Vertical fluxes of heat and momentum are assumed constant. Monin and Obukhov (1954) suggested that the mean wind and temperature gradients through this layer are universal functions of \( z/L \).15 Strictly, MOST applies for stationary, homogeneous turbulence in air over a flat surface with zero mean vertical velocity.

MOST was particularly useful because it was already realized that von Kármán’s (1931) logarithmic wind profile needed correction for nonneutral thermal stratification (e.g., Rossby and Montgomery 1935; Deacon 1953, from 1940 measurements). Many investigators participated in developing the needed functions. Fleagle and Businger’s (1963) text shows a conceptual derivation of the so-called KEYPS16 expression that relates the normalized vertical shear of the mean horizontal wind to thermal stratification:

\[ \phi_m^4 - \gamma \frac{z}{L} A_m^3 = 1, \]  

(9-5)

where \( \phi_m = (kz/u^*) \left( \frac{\partial U}{\partial z} \right) \) and \( \gamma \) is a constant.

12 Though surface energy balance was not the focus, both these papers address the problem of estimating sensible and latent heat fluxes over heterogeneous areas.

13 For a complete summary of M-O similarity theory, see Foken (2006).

14 Through applying the Buckingham Pi theorem to the relevant scales, \( z, u^*, w T, \) and \( g/T \).

15 Apparently, Heinz Lettau (1949) introduced the same length scale, and Joost Businger (1955) independently developed the same similarity theory as part of his Ph. D. thesis, according to his resume, which is available in the NCAR Archives; and confirmed by Businger’s son Steven (November 2018, personal communication). Also see Businger and Yaglom (1971).

16 An acronym based on the names Kazanski, Ellison, Yamamoto, Panofsky, and Sellers, see Panofsky (1963) and Lumley and Panofsky (1964). Businger and Yaglom (1971) suggested expanding to “O’KEYPS” to include Obukhov, since his 1946 paper includes a similar formula.
In the 1960s, (9-5) was modified by Businger and Dyer (Businger et al. 1971), who found from observations for unstable conditions that \( \phi_m \) and \( \phi_h = (kz/\theta_a)(\partial\theta/\partial z) \), where \( \theta_a = \frac{w\partial\theta}{u_a} \), are given by

\[
\phi_m = \left(1 - \gamma_1 \frac{z}{L}\right)^{-1/4}, \quad \text{and} \quad (9-6a)
\]

\[
\phi_h = \phi_{hN} \left(1 - \gamma_2 \frac{z}{L}\right)^{-1/2}, \quad (9-6b)
\]

where \( \gamma_1 \) and \( \gamma_2 \) are constants to be determined, and the subscript \( N \) corresponds to neutral stratification \((z/L = 0)\). The surface-stress estimates from flux plates and mean profile measurements collected near O’Neill, Nebraska, were used to test (9-6a), e.g., Panofsky (1963).

It was the 1968 Kansas experiment (Haugen et al. 1971), during which both eddy-correlation fluxes and profiles were measured, that enabled evaluating flux-profile relationships for both temperature and momentum for stable and unstable thermal stratification. They found that \( \phi_{hN} = 0.74 \), \( \gamma_1 = 15 \), and \( \gamma_2 = 19 \) in (9-6a) and (9-6b) (Businger et al. 1971). For stable conditions \((L > 0)\), they found

\[
\phi_m = 1 + 4.7z/L, \quad \text{and} \quad (9-7a)
\]

\[
\phi_h = 0.74 + 4.7z/L = \phi_{hN}(1 + 6.4z/L), \quad (9-7b)
\]

but there were few points beyond \( z/L = 1 \) on the stable side. Finally, from (9-3a) and (9-3b) and the definition of \( \phi_m \) and \( \phi_h \),

\[
\frac{\phi_h}{\phi_m} = \frac{K_M}{K_H} = \text{Pr}_T,
\]

where the turbulent Prandtl number at neutral stratification \( \text{Pr}_T \) is 0.74.

Subsequent international comparisons of flux measurements revealed inconsistent results, prompting improvements in sonic-anemometer design (Högström 1988) and more sophisticated corrections for flow distortion (Wieringa 1980; Wyngaard 1981). Applying both in a new set of measurements, Högström was able to update the ratio \( \text{Pr}_T \) from its Kansas value of 0.74 to \(~1\), and \( k \) from its Kansas value of 0.35 to 0.4. For \( \text{Pr}_T = 1 \) and \( k = 0.4 \), more recent values of \( \gamma_1 \) are on the order of 16–19.3, and \( \gamma_2 \sim 11 \). The constant 4.7 was revised to be closer to 6 for \( \phi_m \) and 8.2 for \( \phi_h \) (Foken 2006). Refinements have continued for all these variables (e.g., Andreas et al. 2006).

Another important—but avoidable—difficulty in observational testing and evaluation of similarity relationships occurs when nondimensional variables containing shared factors—such as the case for (9-6) and (9-7)—are plotted against one another. This has been called self-correlation. Self-correlation can combine with observations to either mask the physical correlation or augment the physical correlation, a problem that becomes especially serious in the very stable boundary layer (defined in section 5), where small fluxes also make observational errors relatively large. Fortunately, there are methods for assessing self-correlation (e.g., Hicks 1978; Klipp and Mahrt 2004; Baas et al. 2009; Sfyri et al. 2018) or avoiding it altogether (Anderson 2009). Large self-correlation does not necessary negate the utility of similarity theory but it does create an observational ambiguity.

Paulson (1970) integrated (9-6), making it possible to estimate fluxes iteratively, using

\[
\overline{u} = \frac{u_a}{k} \left[ \ln \frac{z}{z_0} - \psi_m \left( \frac{z}{L} \right) + \psi_m \left( \frac{z_0}{L} \right) \right], \quad \text{and} \quad (9-8a)
\]

\[
\overline{v} = \frac{\bar{\theta}}{k} \left[ \ln \frac{z}{z_0} - \psi_h \left( \frac{z}{L} \right) + \psi_h \left( \frac{z_0}{L} \right) \right], \quad (9-8b)
\]

where the roughness length \( z_0 \) is the height at which the mean wind \( \overline{U} = 0 \) and \( \overline{v} = 0 \), and \( \psi_m \) and \( \psi_h \) are corrections for stability. For convenience, (9-8) is written with the integrated \( \psi \) functions evaluated at \( z_0 \) and \( z \) as in Brutsaert (1982), instead of the indefinite-integral form used in Paulson’s (2). Paulson also suggests that \( \phi_h/\phi_{hN} = 1 \) in (9-6b), based on observations at Kerang (Swinbank 1964). Typically, \( z_0 \) for a location is found empirically from (9-8a) for near-neutral conditions; it is larger for rougher surfaces. While it is implicit in (9-8b) that \( \overline{u} = \overline{v} = 0 \) at \( z_0 \), it was recognized earlier that it would be more appropriate to replace \( z_0 \) with \( z_{0m} \) in (9-8a) and replace \( z_0 \) in (9-8b) with \( z_{0h} \) (e.g., Chamberlain 1966).

The relationship of \( z_{0m} \) to \( z_{0m} \) appears to be a function of surface cover (Zilitinkevich 1970; Chen and Zhang 2009) and time of day [Sun (1999), if \( \overline{\theta}(z_{0h}) \) is the surface radiation temperature]. Following Paesckie (1937), \( z \) is replaced with \( z - d \) in the presence of tall vegetation or a canopy, where \( d \) the “displacement height” is the height of an upwardly displaced effective surface. That is, \( d + z_{0m} \) is the zero intercept for a logarithmic wind profile under neutral conditions; data show \( d \sim z_{Ch} \), where \( h_{c} \) is canopy height, and \( C \) is a scaling factor ranging between 0 and 1, with larger values for denser canopies. It follows that \( d \) can be a function of wind direction. Typical values for crops are between 0.6 and 0.7 (Campbell and Norman 1998; Brutsaert 1982).

The exchange coefficients for momentum and heat can be obtained by combining (9-2a) and (9-2b) with (9-8a) and (9-8b) and using the appropriate roughness lengths:
\[ C_D = \frac{k^2}{[\ln\left(\frac{z}{z_{onm}}\right) - \psi_m\left(\frac{z}{L}\right) + \psi_m\left(\frac{z_{onm}}{L}\right)]^2} \] and
\[ C_H = \frac{k^2/Pr_{TN}}{[\ln\left(\frac{z}{z_{onm}}\right) - \psi_m\left(\frac{z}{L}\right) + \psi_m\left(\frac{z_{onm}}{L}\right)]} \left[\ln\left(\frac{z}{z_{onh}}\right) - \psi_h\left(\frac{z}{L}\right) + \psi_h\left(\frac{z_{onh}}{L}\right)\right] \]

The water vapor exchange coefficient \( C_E \) is found analogously, except that \( Pr_{TN} \) is replaced with the Schmidt number for neutral stratification \( Sc_{TN} \). In practice, \( Pr_{TN} \) and \( Sc_{TN} \) are often set to unity. For \( z = 10 \text{ m} \) and \( z_{onm} = 0.1 \text{ m} \) (grass with scattered trees and houses or thick grass), \( C_D \approx 0.01 \) for neutral conditions (\( \phi_m = 0 \)).

The cited work so far assumed the effect of water vapor on the buoyancy flux was negligible. To estimate surface buoyancy flux, it is often assumed that the similarity function for virtual temperature \( \phi_T = \phi_h \). This assumption works, provided (Dias and Brutsaert 1996) that the correlation coefficient between the temperature and mixing-ratio fluctuations \( R_{\theta q} = \frac{\overline{T q f' (\sigma_T \sigma_q)}}{1} \) (1 for stable conditions), although Larsen et al. (2014) argue that MOST relationships could be reliable for \( |R_{\theta q}| \) as low as 0.5. The less stringent criterion appears to be commonly easily satisfied over land, at least to within experimental error for unstable (e.g., Swinbank and Dyer 1967) and stable (e.g., Dias and Brutsaert 1996) conditions. However, measurements over heterogeneous pine forests by Asanuma and Brutsaert (1999) and Lamaud and Irvine (2006) indicated a difference between MOST scalar similarity functions. The latter work inspired Moene and Schüttemeyer (2008) to develop a conceptual model for fluxes over heterogeneous surfaces, with updrafts carrying heterogeneous properties, and downdrafts carrying homogeneous properties. Not only did it reproduce many of Lamaud and Irvine’s results, but it offered a method to estimate scalar fluxes.

Likewise, measurements in the boundary layer over the warm oceans as early as the 1970s revealed a dropoff and even a sign reversal of \( R_{\theta q} \) at longer wavelengths (e.g., Phelps and Pond 1971; Donelan and Miyake 1973), resulting from both a sign reversal in \( w' \theta' \) and an increase in \( \sigma_T \) and \( \sigma_q \) near the surface. This behavior is also related to entrainment as revealed in an LES study by Cancelli et al. (2014). Both contributions are due to “nonlocal” downward heat transport by large eddies extending through a moist mixed layer, where \( \overline{\theta} \) increases with height (section 4). The success of the weaker \( R_{\theta q} \) criterion even over the ocean is consistent with the satisfactory results obtained under the \( \phi_m = \phi_h \) assumption (e.g., Fairall et al. 1996).

For wind, Panofsky et al. (1977) showed using observations that horizontal-velocity variances in the surface layer were a function of \(-z/L\) rather than \(-z'/L\). Even though there are limitations associated with use of LES to evaluate MOST near the lower boundary, results in Khanna and Brasseur (1997) and Johansson et al. (2001) also point to a dependence of horizontal velocity variances on \(-z'/L\), with the latter paper suggesting a weak dependence of temperature variance as well. Finally, while Katul et al. (2011, 2013) have put \( \phi_h \) and \( \phi_m \) on a more rigorous basis, they have also offered insight into their limitations (because of the impact of large eddies).

c. Vegetation canopies

When the study of plant canopy turbulence was in its infancy, researchers assumed that it could be treated like boundary layer turbulence with the addition of finescale eddies generated in the wakes of the leaves, twigs, and branches. By the 1970s, it had become clear that the dominant turbulent eddies in plant canopies are much larger than plant element size. However, it took two decades to show how these eddies are generated, and another decade to develop methods to account for the associated fluxes in parameterization schemes.

1) Statistics

Mean profiles within a canopy (shown for neutral conditions in Fig. 9-3) are distinct from profiles in the atmospheric surface layer. The mean streamwise velocity profile follows a standard logarithmic form well above the canopy, while the profile within the canopy can be roughly described as exponential (e.g., Cionco 1965). The two types of profiles merge within the roughness sublayer (RLS) and an inflection point appears at the top of all but the sparsest canopies. The size of the most energetic eddies increases with height in the surface layer (Kaimal and Finnigan 1994). However, as will be shown below, this height dependence vanishes with descent through the...
roughness sublayer, and the size of the dominant canopy eddies is height independent.

Above the canopy, momentum fluxes are nearly constant with height but rapidly decay with descent into the canopy as streamwise momentum is absorbed through canopy drag; for dense canopies $-\overline{\nu w}$ reduces nearly to zero by ground surface, indicating that nearly all downward transport of horizontal momentum has been absorbed by the canopy and therefore does not reach the underlying surface. Hence, within the canopy, the similarity scaling leading to the log-law and MOST no longer applies and the direct connection between local gradients and turbulent fluxes of scalars and momentum is lost (e.g., Denmead and Bradley 1985; Finnigan 1985), and turbulence in the RSL transports momentum and scalars more effectively than it does in the surface layer above. Turbulent momentum transport near the canopy top is by frequent upward motions bringing low-momentum air upward (Q2, ejections), with infrequent but stronger downward penetrations of high-momentum air (Q4, sweeps). Though less frequent, the sweeps contribute more to the exchange of momentum and scalars in the RSL than do the ejections (Finnigan 1979; Shaw et al. 1983; Katul et al. 1997; Finnigan et al. 2009). This is in contrast to surface-layer flows, where the contributions from sweeps and ejections are of comparable magnitude or in which ejections dominate (Raupach and Thom 1981).

In the surface layer well above the canopy, the budget of TKE [$= 1/2(\overline{u'^2} + \overline{v'^2} + \overline{w'^2})$, where $\nu$ is the horizontal wind component normal to $\overline{u}$] is generally in local equilibrium, with shear production balancing dissipation at any height under neutral conditions. However, within and just above a canopy, TKE produced in the region of high streamwise velocity shear near canopy top is exported by pressure and turbulent transport to regions away from its source. Hence at any given height within the RSL, the local rates of production and dissipation are not usually in balance, and local equilibrium is lost. All canopy second-order budgets\(^{18}\) lack a local balance between production and dissipation, because the dominant turbulent eddies in the canopy span most of its depth and thus can efficiently transport TKE and scalars through the entire canopy. Furthermore, the thin viscous boundary layers on the surfaces of all the canopy elements generate sufficiently abundant shear layers at the Kolmogorov microscale that the dissipation rate of turbulence within canopies can be several times higher than seen in free air flows with shear of comparable magnitude.

In canopy flows, assumptions in Kolmogorov's (1941b, 1962) theory are violated in a number of ways. First, interactions with the canopy elements act to convert mean kinetic energy into turbulent kinetic energy at scales tied to the leaves, twigs and branches; these scales usually fall at the high wavenumber ($k$) end of the inertial subrange. This spectral-shortcut process takes TKE from lower wavenumbers and injects it directly into high wave-numbers. Additionally, the work eddies do against the viscous drag of the canopy removes energy at all scales,\

\(^{18}\) E.g., of $\overline{s^2}$, $\overline{w'i}$, and $\overline{w'i'}$, where $s$ is a scalar and $i = 1, 2, 3$ indicate the three wind components.
including inertial subrange scales. This latter process changes the overall magnitude of canopy dissipation. The net result is that within canopies $E(k)$ in the inertial subrange decreases faster than $k^{-5/3}$ at low wavenumbers and more slowly at high wavenumbers. For further discussion, see Finnigan (2000).

2) ORGANIZED CANOPY-INDUCED TURBULENT MOTIONS

Numerous observational studies have attempted to deduce the three-dimensional spatial flow structures underlying the observed scalar ramp patterns (e.g., Gao et al. 1989; Zhang et al. 1992; Shaw et al. 1995; Finnigan and Shaw 2000); however, the difficulty in making the necessary observations prompted researchers to turn to LES (see section 2c). Following the pioneering studies by Shaw and Schumann (1992) and Fitzmaurice et al. (2004), Finnigan et al. (2009) used LES to deduce the three-dimensional structure by referencing fields to the canopy-top pressure maxima to ensemble average the flow field. They found that in neutrally stratified flow, the mean organized structure consists of a combination of “head-up” and “head-down” hairpin vortices, which are both inclined in the downstream direction (Fig. 9-5). The ejections and sweeps are flows induced between the counterrotating legs of the hairpins. The convergence region between the downstream ejection and upstream sweep produces a scalar microfront and a positive static pressure pulse.

The structures in Fig. 9-5 are consistent with theory, as described in Patton and Finnigan (2013). First, drawing on the analogy between the canopy-top inflection point in $U$ and that which develops as a result of mixing at the interface between two free-atmosphere layers with different wind speeds, Raupach et al. (1996) shows that the inviscid instability at the inflection point under neutral stratification leads to the development of
Kelvin–Helmholz (K-H) waves. These in turn roll up into “cat’s eye” vortex tubes (Stuart 1967). Finnigan et al. (2009) suggest that adjacent vortex tubes, in response to random turbulent perturbations, approach and rotate around each other and distort into hairpins in a process matching the helical-pairing instability of Pierrehumbert and Widnall (1982), and finally align themselves with the shear, as shown in Fig. 9-5. It should be noted that Bailey and Stoll (2016) produced a somewhat different eddy structure from a canopy LES, at least partially because their analysis incudes the total streamwise velocity rather than its perturbation.

3) ROUGHNESS SUBLAYER PARAMETERIZATION

Although some details of Finnigan et al.’s (2009) theory remain speculative, it successfully explains most observations in neutrally stratified flows. Indirect evidence for the fundamental role played by this canopy-top instability comes from Harman and Finnigan’s (2007, 2008, hereafter HF) extension of MOST to the RSL, which uses the vorticity-layer thickness \( \delta_m = \frac{2(\bar{u})(\partial \bar{u})/\partial z) \) evaluated at canopy top, where the angle brackets represent a horizontal average, Raupach et al. (1996) associated with the K-H waves at canopy top as an additional scaling length to modify the MOST streamwise velocity (\( \bar{U} \)) profiles. The HF equation contains two sets of terms. The first is analogous to (9-8) but is rewritten such that the lower surface is at canopy height \( h_c \) rather than at the apparent sink for momentum \( (d + z_{0m}) \); rewriting the equation in this way eliminates \( z_{0m} \). The second set of terms accounts for canopy-induced physics in the RSL. Key aspects of the HF parameterization include 1) \( d \) and \( z_{0m} \) are flow-dependent parameters rather than specified based on canopy structure, and 2) the formulation naturally relaxes back to standard MOST above the RSL.

The HF parameterization reproduces observed mean velocity and scalar profiles through the RSL for a wide range of canopy types and stabilities. Weligepolage et al. (2012) tested numerous canopy parameterizations against observational data collected over a fir forest and found that HF performed best. More recently, Bonan et al. (2018) implemented the HF formulation within the Community Land Model (Oleson et al. 2010) and found substantially improved skill in predicting surface exchange at a number of forest, cropland, and grassland sites that span a broad range of canopy height, leaf area index, and climate.

4) FUTURE DIRECTIONS

Although the theory proposed by Raupach et al. (1996) and Finnigan et al. (2009) and the parameterization based on that theory by HF appear to work quite well over a range of canopies and stabilities (e.g., Bonan et al. 2018), the theory relies upon the presence of mean velocity shear at canopy top. However, under strongly stratified conditions (unstable and stable), mean shear largely vanishes. Hence, the parameterization’s applicability at stabilities far from neutral becomes suspect.

Diabatic influences on canopy-turbulence statistics are reasonably well established (e.g., Shaw et al. 1988; Fazu and Schwerdtfeger 1989; Leclerc et al. 1990; Mahrt et al. 2000; Cava et al. 2004). Under unstable conditions, turbulent scalar exchange remains vigorous even in periods with weak mean shear at canopy top (e.g., Katul et al. 1997; Dupont and Patton 2012a; Patton et al. 2016). Under strongly stable conditions, buoyancy damps vertical motions (e.g., Lee 2000; Cava et al. 2004; Vickers and Thomas 2014), and the canopy layer can decouple from the overlying ASL (e.g., Belcher et al. 2012; Thomas et al. 2013), resulting from differing efficiencies of heat versus momentum transport. While progress has been made toward characterizing buoyancy’s influence on organized canopy-scale structures (e.g., Gao et al. 1989; Cava and Katul 2008; Thomas and Foken 2007; Dupont and Patton 2012b; Shaw et al. 2013), much work remains to elucidate the turbulent eddy structure performing momentum, energy, and scalar exchange under nonneutral conditions.

In addition, this discussion has assumed that the canopy consists of horizontally homogeneously distributed elements located on flat terrain; however, these conditions are rarely found outdoors. Rather, heterogeneously distributed vegetation is the norm, whether as a result of human action through wind breaks, clear cuts, forest edges, transitions from one crop to another, etc., or through natural transitions in surface cover. Currently, there is no universal theory describing how canopy flow reacts to stratification, heterogeneity, or topography although knowledge is accumulating in incremental steps.

d. The surface layer over the ocean

1) PROBLEMS UNIQUE TO THE OCEAN SURFACE

Because the ocean is a fluid, near-surface atmospheric measurements are difficult and the physics challenging. Turbulent fluxes are a challenge to measure from a
moving platform that is sprayed with saltwater in stronger winds. Estimating the mean quantities for the bulk aerodynamic technique (9-2) is also more difficult than over land. Since the ocean surface moves, \( U \) in (9-2) must be relative to the surface motion, which derives not only from the great ocean currents such as the Gulf Stream, but also from the superposed Stokes drift (resulting from the net motion of the orbiting water parcels in the direction of wave propagation), and the drift current (resulting from the drag of the air on the ocean surface). Likewise, the sea surface temperature (SST) typically has been measured not at the air–sea interface, but rather at a ship’s intake or at some depth below a buoy or platform; this is called the bulk SST. Finally, there are waves. Waves driven by the wind (so-called wind sea) are superposed on swells from any direction. In high winds, the waves not only grow in amplitude but they also give off spray. Some or all of these phenomena affect the surface fluxes in interesting ways, many the subject of active research.

2) Measurement methods—A brief history

(i) Bulk methods

Estimation of surface fluxes was pioneered by Montgomery using bulk methods (for more detail see Sverdrup et al. 1942), with further work by many others. The earliest estimates of sensible heat flux, evaporation, and wind stress were from research ships such as the German Research Vessel (R/V) Meteor (Defant 1932; Kuhlbrodt and Reger 1938). It was quickly found that the drag coefficient \( C_D \) was not a constant but increased with the 10-m wind speed \( U_{10} \) (Montgomery 1940; Deacon et al. 1956).

Most bulk formulations have used bulk SST rather than the actual or interfacial SST, which was not available; adjustments depend on the method of measurement and the formulation used. Today, infrared (IR) and microwave sensors (on ships, aircraft, or satellites) measure the outgoing radiation from a layer just below the interface (less than 1 mm for IR, slightly farther down for microwaves) to estimate the interfacial SST; but corrections are required for atmospheric transmission, reflections of the atmospheric and cloud radiation off the sea surface, and possibly sun glint (Robinson 2004). When using the interfacial temperature, consideration should be taken of the so-called cool film—at least at low wind speeds, temperatures can be up to 0.2–0.3 K lower at the interface than a few millimeters below because of loss of heat from the water to the atmosphere. However, at low wind speeds and under clear, sunny skies, a warm layer develops in the upper ocean because of absorption of sunlight and weak mixing, and the bulk and infrared surface temperatures can differ substantially. To relate these temperatures requires keeping track of the conditions over a period of time and preferably coupled modeling of the upper ocean and atmosphere [for reviews, see Katsaros (1980) and Fairall et al. (1996)]. Much research on the long-term bulk SST and the interfacial SST continues, because of the importance for coupled models and climate change evaluations (e.g., Donlon et al. 2007; Kent et al. 2017).

(ii) Eddy correlation

Today the preferred method to obtain fluxes in situ is by eddy correlation, which requires a stable platform or corrections for the platform motion in three dimensions. Turbulent fluxes collected by Liu et al. (1979) from a bottom-mounted mast over Lake Washington (Seattle, Washington) using this technique resulted in formulations of the exchange coefficients over water that have been used widely and extended to higher wind speeds by several field campaigns (e.g., Fairall et al. 1996, 2003; Edson et al. 2013). In addition, the “dissipation method,” which equates the turbulent dissipation at small scales with the production by eddies at larger scales, has been used (e.g., Large and Pond 1981). These methods have been compared by Donelan et al. (1997), who report good agreement for wind sea dominated conditions, but less well in the presence of swell. Turbulent fluxes have also been obtained from moored buoys, with the preferred instrumentation first being pressure sensors or propeller systems for the three-dimensional wind components (Smith 1980; Large and Pond 1981). Developments in the 1990s reduced wetting and salt-accumulation problems, allowing sonic anemometers to be operated unattended for long periods, even on autonomous buoys (Fairall et al. 1997; Edson 2009). Sonics provide reasonable measurements of sensible heat flux, but water vapor flux still suffers from contamination of humidity sensors by rain and sea spray.

Since the 1960s, the Floating Instrument Platform (FLIP), a ship designed to rotate 90° from its normal position so that its bow is pointed upward, has allowed eddy correlation and other turbulence, profile, and oceanographic measurements with minimal effect from ship motion; enabling studies of sea–air interaction over the open ocean (e.g., Pond et al. 1971; Rieder et al. 1994; Hristov et al. 2003; Zappa et al. 2012). Several other stable floating platforms, extended piers away from shore, and bottom-mounted research stations in relatively shallow water have also been used extensively. An early example used successfully was the floating mast (anchored to the surface 59 m below) off Nova Scotia used by Smith (1980) to obtain wind stress by the eddy correlation method.
(iii) Scatterometry and importance of shorter waves to stress

Surface stress and the $\overline{U}$ profile are expected to be related to the mean-square slope of the sea surface, which can be determined from the spread of the sun glint off the water surface, as observed from aircraft by Cox and Munk (1954a, b). One of the early suggestions how the roughened sea surface affects the wind stress was by Charnock (1955), who formulated a simple relationship between the roughness of the sea to waves based on dimensional grounds:

$$z_{um} = \frac{u^2_a}{\alpha}, \quad (9-10)$$

where $\alpha$ is the Charnock “constant,” whose value has been estimated in many publications and turns out to vary (e.g., Komen et al. 1998). This relationship led to the development of radars (scatterometers) to remotely sense the surface wind and stress over the ocean from space.

Scatterometer measurements suggest an important role of the small capillary/gravity waves in determining stress. Scatterometers receive the so-called Bragg scatter (Bragg-scattered radiation) from waves of wavelength commensurate with the radar wavelength, which was 2.1 cm (14.6 GHz) on the first scatterometer flown on the Seasat satellite in 1978. Other wavelengths were used on later instruments (e.g., Robinson 2004). Since the scatterometer reacts to the sea surface roughness at small scale, it is a measure of the wind stress $u_a$. The strength of the Bragg scatter was typically calibrated against the wind speed at 10-m height, measured from buoys and corrected to neutral stratification. This relationship is now well established for wind speeds up to about 20 m s$^{-1}$. Currently, it is generally thought that waves of length smaller than $\sim 10$ m are responsible for the bulk of the drag at winds of about 10 m s$^{-1}$. This is a key concept since a typical wave spectrum contains scales as large as 250 m or more. However, it is still not completely settled as to whether the few intermittent big breakers associated with longer-wavelength waves could be as important as many small microbreakers in transmitting stress from the atmosphere to the ocean, particularly in strong winds.

3) Fluxes and the impact of waves

Many aspects of the wave field—such as phase speed, slope, and peak amplitude (Kitaigorodskii 1959) and interactions across the wave spectrum (Kitaigorodskii 1983, Hasselmann et al. 1973)—were considered as contributors to the wave drag on the atmosphere and generating turbulence in the ocean. This has led to increasing interest in the sea state as well as the surface stress, but typically the 1D frequency spectrum of wave height was obtained but not the directional wave spectrum. Geernaert et al. (1986) reported the wind stress under conditions of fully developed wind sea for measurements at the Nordsee platform with steady winds and a long fetch; Donelan et al. (1997) reported the results of the Surface Wave Dynamics Experiment (SWADE), which was aimed at measuring the effects of waves on drag. The directional wave spectrum including swell was measured during SWADE, but not the high frequency, capillary, and gravity waves (Drennan et al. 1994).

More recently, field measurements have clarified the effect of moving waves on near-surface turbulence. The OHATS (section 2c) and the Coupled Boundary Layers Air–Sea Transfer (CBLAST; Edson et al. 2007) measured the impact of fast-moving swell on air motion, scalars, and turbulence. These results clearly showed how the angle between wave motion and relative wind influences low-level turbulence and profiles, to the point that $C_D < 0$ under conditions of weak wind in the presence of swell. Sullivan et al. (2008) combine idealized LES simulations of the interaction between air and moving waves to CBLAST observations to isolate the mechanisms involved. These authors point out that hints of this unusual behavior can be traced back to Harris (1966), who found wave-driven wind in wave-tank experiments; and cite subsequent observational studies that have associated waves with unusual near-surface behavior of the atmosphere, such as unusual wind profiles and momentum fluxes not aligned with the mean wind, as well as upward momentum flux.

Recently, more focus has been on extremely high winds. For a wind sea adjusted to the wind speed, Edson et al. (2013), in presenting the latest version of the Coupled Ocean–Atmosphere Response Experiment (COARE) flux algorithm (version 3.5), suggest that wave influences need not be modeled separately. This is because both the influence of the wave field and the shear-induced drag are strongly correlated with the wind speed, so that both effects can be included in the formulation of the dependence of $C_D$ on wind speed. Figure 9-6 illustrates this relationship and shows the recent estimates of $C_D$ as well as the formulation used by the coupled model at the European Centre for Medium-Range Weather Forecasts (ECMWF).

However, the individual $C_D$ estimates at high wind speed show wide scatter, and there are some situations when the wave field is not adjusted to the wind, such as in tropical cyclones (TCs). In a moving TC, where swell and old wind sea (wind sea in equilibrium with the forcing wind) are not aligned with the rotating wind around the storm center, much more complex relationships occur. Chen et al. (2013) have modeled these interactions for the
different sections of a moving TC, showing that relationships are asymmetric between the right- and left-hand portions relative to its direction of motion.

Measurements of heat and moisture fluxes have been more difficult because of contamination from soot on ships, rain, and sea spray. Thus, acceptable values have mostly been determined under benign conditions at lower wind speeds. Based on the earliest measurements (e.g., Liu et al. 1979), the exchange coefficients $C_H$ and $C_E$ are constant for $U_{10} < \sim 10 \text{m s}^{-1}$. The Humidity Exchange over the Sea (HEXOS) field program extended this result to $U_{10} = 18 \text{m s}^{-1}$ (DeCosmo et al. 1996). The measurements used special techniques to eliminate the sea spray from the airstream before taking measurements of temperature and humidity fluctuations (Katsaros et al. 1994). Though no profiles were obtained under spray conditions, single-height $q$ measurements made it clear that the reduction in surface evaporation compensated “mostly” for the addition of humidity from the evaporating spray. Later measurements of turbulent fluxes of water vapor on the outskirts of hurricanes from aircraft during CBLAST by Zhang et al. (2008) confirmed the HEXOS results and found the exchange coefficients for heat and water vapor to be constant to wind speeds as high as 25 m s$^{-1}$ (Fig. 9-7), albeit with large scatter. This implies that in most conditions, when wind sea is in equilibrium with the wind, the effect of the waves might be incorporated in the bulk formulation for sensible heat and water vapor flux. The physics of how surface waves modify the turbulent transfer of sensible heat and water vapor is, however, still not well established.

Work continues on fluxes at extremely high winds. The effects of sea spray, though the subject of much speculation and study but few successful measurements, are still not well understood, according to a recent review by Veron (2015), though Andreas et al. (2015) used a bulk flux algorithm to suggest some effects at high wind speeds; while Richter and Sullivan (2014) and Helgans and Richter (2016) use DNS to better understand the mechanisms involved and their net impact on heat and moisture fluxes. Other promising work at extreme wind speeds has included laboratory studies (e.g., Donelan et al. 2012; Haus et al. 2010), LES (e.g., Sullivan et al. 2014) and even a budget study to back out the exchange coefficients in two hurricanes (Bell et al. 2012). Such methods may be the only viable means to

FIG. 9-6. A compilation of drag coefficient values as a function of wind speed after a reevaluation of many field programs, labeled COARE 3.5 (representing the COARE algorithm, version 3.5), where the data have been averaged over wind speed bins; it includes a curve for the case including wave age dependence and an illustration of the drag formulation used in the forecast model at ECMWF. [Figure adapted from Edson et al. (2013)].

FIG. 9-7. The exchange coefficient for evaporation with results from several field experiments. The experiments all used the eddy correlation method to measure the fluxes. Non-CBLAST data are from AGILE (Donelan and Drennan 1995), HEXOS (DeCosmo et al. 1996), GASEX (McGillis et al. 2004), SOWEX (Banner et al. 1999), and SWADE (Katsaros et al. 1993). Data were corrected for density variations (Webb correction) and salinity effects using methods of Fairall et al. (2003). [Figure adapted from Drennan et al. (2007)].
estimate exchange coefficients in extreme winds ($U_{10} > 30 \text{ m s}^{-1}$), but they show large scatter of the resulting exchange coefficients and thus need verification against in situ data to wind speeds as high as can be managed. For a recent review of the coupling between surface waves and winds and currents under broad range of conditions, see Sullivan and McWilliams (2010).

4. CBL

a. Mean profiles

By the early twentieth century, dry-adiabatic lapse rates were regularly observed in daytime summer U.S. Weather Bureau (USWB) kite soundings over the center of the country, even in monthly averages (e.g., USWB 1918). As for the corresponding wind profile, many students in the midlatitudes were taught up until the 1970s that it would be similar to the “Leipzig wind profile” (Mildner 1932), whose hodograph somewhat resembled the classic Ekman (1905) spiral. In fact, even the CBL depth was deemed proportional to the Ekman layer depth scale $u_{l0}/f$ (e.g., Tennekes 1970; Kaimal et al. 1976), even though tropical meteorologists were routinely using jumps in $\frac{\partial q}{\partial z}$ and $\frac{\partial q}{\partial z}$ to define boundary layer top decades earlier (e.g., Bunker et al. 1949, section 8).

While Ekman’s spiral assumed a constant $K_M$, the Leipzig profile $K_M$ varied with height, reaching a maximum in the lower ABL (Lettau 1950). Following Prandtl (1932), Blackadar (1962) used $K_M = \frac{S}{l}$, where $S$ is the magnitude of the vector shear and $l$ is a mixing length (or characteristic turbulence length scale) that depends on height along with a characteristic eddy size $\lambda$ (~150–250 m), to produce wind profiles whose turning angle increases with surface roughness. This length, given by the expression

$$l = k\frac{z}{(1 + \frac{kz}{\lambda})},$$

(9-11)
is proportional to $z$ near the surface, as postulated by Prandtl (1925) and equal to $\lambda$ in the free atmosphere ($kz/\lambda \gg 1$). This relationship is still used in ABL parameterization schemes (e.g., Hong et al. 2006).

The turning of the wind with height can be explained in terms of force vectors (drag, pressure gradient force, Coriolis force) at the surface and the top of the ABL (pressure gradient force and Coriolis force) (Fig. 9-8) or mathematically by

---

20 Even though the Leipzig wind profile was based on a set of pilot-balloon observations collected under stable conditions, according to Lettau (1950).
where $U$ and $V$ are orthogonal horizontal wind components, the subscript $g$ indicates geostrophic wind, and $f$ is the Coriolis acceleration. Most observed wind profiles were not so well behaved: while veering with height was commonly observed in Northern Hemisphere soundings, the hodographs tended to be quite irregular. At about the same time, Faller (1963) was finding that Ekman flows in the laboratory broke down into longitudinal rolls, prompting Faller and Kaylor (1966) and other investigators (e.g., Lilly 1966b; Brown 1970) to investigate the stability of the Ekman spiral. They found instabilities that they associated with the origin of longitudinal rolls. Brown allowed his most unstable mode to grow to equilibrium (net exchange of energy with the mean flow) to produce a strongly modified hodograph with much less veering. Brown (1978) matched his Ekman-instability solutions to the surface layer to produce increased turning angle with roughness (surface drag) and stable thermal stratification.

The advent of LES in the early 1970s (see section 2c) led to near-universal acceptance of thermodynamic variables as the primarily indicators of CBL depth (e.g., Deardorff 1970a, b, 1974a), and along with new measurement capabilities, an increased emphasis on the whole ABL. Our current picture of the fair-weather CBL appears in Fig. 9-9. In the figure, the height is normalized by $z_i$, the top of the ABL as defined by the height of the minimum buoyancy flux $B = (g/\omega_f)\bar{w}^' \theta^'$ (see Fig. 9-10). This definition is not practical for the observed ABL, so the height of the mixed-layer top $h_1$, where $\bar{\theta}$ starts to increase with height (Fig. 9-9c), is often used. In Fig. 9-9, the surface layer corresponds to where $\bar{\theta}$ and the longitudinal and lateral wind components $U$ and $V$ change with height, the mixed layer has near-constant values, and the transition layer lies between $h_1$ and $h_2$, where $\bar{\theta}$, $U$, and $V$ reach their free-atmosphere values. While the shear-driven ABL has an Ekman-like profile, the CBL has almost no wind turning with height, except near the surface and above $h_1$. This well-mixed CBL picture was reinforced by observations in the 1973 Minnesota (United States) experiment (Kaimal et al. 1976) and the observations at Wangara (Australia; Clarke et al. 1971). However, other observations (e.g., LeMone et al. 1999) and more recently LES (Conzemius and Fedorovich 2006a) reveal that these well-mixed layers are features of more slowly growing ABLs with little shear above the ABL.

b. Fluxes

1) BUOYANCY AND MOMENTUM FLUX

The vertical fluxes of buoyancy and horizontal momentum for Moeng and Sullivan’s (1994) convective ABL are plotted as a function of height in Fig. 9-10.
In the figure, the buoyancy flux $B$ is normalized by $w_0/\bar{z}_i$, where the Deardorff (1970a,b) convective velocity scale $w_0 = [(g/\bar{z}_i)(\bar{w}'T_0)]^{1/3}$ (appropriate since $q = 0$ in their simulation but now written using $T_0$). Note that $B$ is nearly linear with height up to $\bar{z}_i$, consistent with the time rate of change of temperature being nearly independent of height but not consistent with $\partial \bar{u}/\partial z$. Such “nonlocal” fluxes\(^\text{22}\) were anticipated by Priestley and Swinbank (1947); and countergradient heat fluxes in the mixed layer were documented from aircraft data by Bunker (1956) and Telford and Warner (1964). From (9-12), the near-linear variation of $\bar{u}'w'$ with height is consistent with near-constant $\bar{V} - \bar{V}_r$ and $\bar{U}$ within the mixed layer, and, thus, nonlocal fluxes.

Figure 9-11 shows variances of the horizontal- and vertical-velocity components normalized by $w_0^2$. In the figure, $\bar{w}'^2/w_0^2$ peaks between 0.3 [Air Mass Transformation Experiment (AMTEX)] and around 0.5 (Willis and Deardorff 1974); the peak in Moeng and Sullivan (1994, not shown) is around 0.4. The variance in the horizontal wind components peaks at the surface, with a secondary maximum near ABL top in both cases. Horizontal-velocity variances for Willis and Deardorff’s (1974) laboratory model follow the same pattern. Temperature and humidity variance (not shown) also have peaks near the top and bottom of the ABL, where $|\partial \bar{\theta}/\partial z|$ are largest; but often their peak near $z_i$ is greater than that at the surface.

2) THE TKE BUDGET AND ENTRAINMENT

Since the 1970s, we have been able to use LES and observations to evaluate terms in the TKE budget, given by

$$\frac{\partial E}{\partial t} = \frac{g}{\bar{	heta}} \left[ u'w' \frac{\partial \bar{U}}{\partial z} - v'w' \frac{\partial \bar{V}}{\partial z} - \epsilon \right] - \frac{\partial (\bar{w}'^2)}{\partial z} - \frac{1}{\rho} \frac{\partial (\bar{w}' \bar{p}')}{\partial z}. \tag{9-13}$$

where $\epsilon = \text{TKE}$, and the energy-production terms on the right side are respectively for buoyancy $B$ and shear $S$ (2 terms), and dissipation $D$. The last two terms represent vertical divergence of vertical transport of TKE $T$ and pressure fluctuations $P$, respectively. Building on the work of Telford and Warner (1964), Lenschow (1970) published the first TKE budget derived from aircraft gust-probe data, which bears a strong resemblance to that for Moeng and Sullivan’s (1994) buoyancy-driven ABL (Fig. 9-12), despite his use of temperature rather than virtual temperature (implying dry air), and statistical uncertainty in the aircraft averages (see, e.g., Mann and Lenschow 1994).

A considerable amount of research has been dedicated to estimating the buoyancy flux at the BL top $z_i (\bar{B}_{z_i})$, where the turbulent boundary layer is capturing (entraining) air from above. An entrainment ratio

---

\(^\text{22}\) Nonlocal fluxes are in contrast to those proportional to $K$ times the local vertical gradient, such as those specified for the surface layer in section 3. Also called nonlocal transport. For further discussion see Deardorff (1966).
$A_B = -B_z/B_0$, where $B_0$ is the surface buoyancy flux, is typically used.\textsuperscript{23} However, since the shape of the area of negative buoyancy can vary, Conzemius and Fedorovich (2006a) use areas as well as $A_B$. A simple interpretation is that the “negative area” for $B$ in Fig. 9-10—that area between the zero axis and the negative buoyancy fluxes, exactly balances the remaining TKE production/destruction within the CBL. Thus, for a CBL with no shear, the “negative area” would equal the integrated contribution of positive buoyancy to the TKE, minus the energy lost to dissipation. The first to estimate this ratio was Ball (1960), who found $A_B = 1$ assuming no shear or dissipation. With time, researchers converged on a ratio for a near-steady-state, weak-shear CBL close to 0.2 (e.g., Stull 1976). However, the negative area becomes smeared out vertically for rapidly growing CBLs, as can be seen by buoyancy profiles idealized using Fig. 9-10a as a function of height for different times, leading to smaller values of $A_B$ (Fig. 9-10c). One can see this smearing effect for more rapidly growing boundary layers in the classical laboratory studies of Deardorff et al. (1969).\textsuperscript{24} Such effects, which provide strong incentive for simulating near steady-state ABLs, can mislead those doing observational studies of evolving ABLs, unless they effectively normalize by focusing on changes in the mixed layer or the entrainment layer (e.g., Betts and Ball 1994; Barr and Betts 1997). For stronger winds [e.g., the Moeng and Sullivan (1994) shear + buoyancy ABL], the extra TKE supplied by shear production allows larger values of $A_B$, something confirmed by observations from the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig9-11}
\caption{Variances of fluctuating horizontal- and vertical-velocity components from the Air-Mass Transformation Experiment [AMTEX; black, from Lenschow et al. (1980)], with $u$ in the direction of the mean wind, and from Willis and Deardorff (1974, Fig. 5) laboratory experiments (red). The different symbols are for different days/ experiments. Idealized curves are based on the Minnesota experiment (Kaimal et al. 1976) and AMTEX (right panel). For comparison to LES see Moeng and Sullivan (1994).}
\end{figure}

\textsuperscript{23} If $B$ is used to define the entrainment ratio, use of virtual temperature or height-corrected virtual temperature will give the same answer. If, however, $\tilde{w}\, \tilde{\theta}_v$ is used, where $\tilde{\theta}_v$ is virtual potential temperature, $A_B$ will be too large by a factor of roughly $[1000/ P(z_i)]^{0.287}$, roughly 3\% for pressure $P(z_i) = 900$ hPa for a $P_0 = 1000$ hPa.

\textsuperscript{24} The more slowly growing boundary layers in Deardorff et al. (1969) have $A_B$ closer to 0.2. According to Jonker (2018, personal communication), this turned out to be because—whether through insight or luck—their laboratory setup was in the right part of the parameter space, as defined by Prandtl, Reynolds, and Peclet numbers (the last being the heat-transfer analog to the Reynolds number). For engaging discussions of the influence of laboratory setups on entrainment results, see Jonker et al. (2012) and Jonker and Jiménez (2014). We also note that LES that use dynamic models to estimate the subgrid-scale turbulent Prandtl and Schmidt numbers are most successful in reproducing realistic results (Shi et al. 2018a). In this case, $Pr_t \approx 10$ in the cloud layer and $-0.5$ in the air beneath.
First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE; Betts and Barr 1996; Barr and Betts 1997). Conzemius and Fedorovich (2006b) summarize some entrainment parameterizations for ABLs with significant vertical shear.

c. Humidity and the importance of virtual temperature

So far, we have dealt with “dry” CBLs. Often, even cloudless CBLs are moist; this means that $\bar{\theta}_w$ is nearly constant in the mixed layer, with $\bar{\theta}$ increasing with height and $\bar{q}$ decreasing with height. Such behavior is especially common over warm oceans (e.g., Bunker et al. 1949). Here, where sea–air temperature differences are often far less than over land, latent heat flux $Q_E$ is correspondingly much larger than sensible heat flux $Q_H$. A consequence is that while $\bar{w}T_y$ remains positive through much of the mixed layer, $\bar{w}T$ is mostly negative, compensated for by large and positive $\bar{w}q$. Negative $\bar{w}T$ was observed early on. As noted in her article in The Sea, Malkus (1962, p. 210) found that temperature anomalies in updrafts went from positive to negative a few hundred meters above the tropical ocean; but the expected positive humidity fluctuation needed to provide buoyancy did not seem to be there—probably because the water vapor sensor was too slow.25 Not long after those observations were taken, both a wet-bulb instrument (Telford and Warner 1962, 1964) and microwave refractometer (McGavin and Vetter 1963; Bean et al. 1972) made fast water vapor measurements and thus $\bar{w}q$ measurements possible.

Examples of $\bar{w}T$, $\bar{w}q$, and $\bar{w}T_y$ profiles collected over the tropical Atlantic Ocean during GATE (1974) in Fig. 9-13 show the largest negative $\bar{w}T$ values corresponding to the largest upward $\bar{w}q$ at the top of the mixed layer, which is just below cloud base. The differences between the flux profiles (based on measurements on roughly a $30 \times 30$-km scale) are related to the amount and vigor of cumulus cloud overhead. These differences appear to average out over larger scales, since budget studies over fair-weather domains on the order of hundreds of kilometers [Riehl et al. (1951) (see section 8c); Augstein et al. (1973) from the Atlantic Trade-wind Experiment (ATEX); Holland and Rasmusson (1973) and Esbensen (1975) from BOMEX] all find cloud-base $Q_E$ on the order of 80%–90% of its surface value.

25 This reinforced the erroneous idea that marine tropical cumulus do not have “roots,” based on aircraft subcloud turbulence observations in Bunker et al. (1949), which persisted until fast water vapor instruments were available.
If we normalize the $w_0T_0$, $w_0q_0$, and $w_0T_0y_0$ profiles by their surface values and combine the resulting profiles for the four days in Fig. 9-13, a linear fit yields a mixed-layer top flux about $2^{0.17}$ times the surface value. While there is a strong relationship between clouds and the $Q_E$ profiles, the cloud base acts much like the top of a cloud-free CBL.

d. Structure

Data from the Minnesota experiment (Kaimal et al. 1976) show that the peak in size for $u$ and $v$ eddies, though broad, is roughly constant through the ABL (Fig. 9-14) with a horizontal wavelength of about 1.5$z_i$. That is, the fluctuations of the surface winds are largely determined by large eddies extending through the CBL. In contrast, $w$ scales increase with height up to about 0.1$z_i$, a consequence of updrafts associated with smaller-scale convergence and mass continuity. At smaller scales, points lie along the $-2/3$ slope associated with the frequency-weighted inertial subrange.

What do the eddies look like? As described by Woodcock (1941), circular soaring of gulls reflected columns of rising buoyant air (cells) in weak winds, while linear soaring in stronger winds reflected along-wind linear updrafts associated with helical circulations (longitudinal rolls). Kuettner (1959) suggested that cloud streets lie above roll updrafts. Hardy and Ottersten (1969) were able to document clear-air rolls and cells with radar. In the 1960s and 1970s, satellite images and photographs by astronauts were showing cloud streets to be common, but some researchers remained skeptical about their relationship to boundary layer motions despite roll-like instabilities in boundary layers with Ekman wind profiles (e.g., Faller and Kaylor 1966; Brown 1970). This prompted LeMone and Pennell (1976) to document an example of direct association of lines of small cumulus with upwelling regions (Fig. 9-15). Weckwerth et al. (1997) used correlation functions from radar returns to determine the type of boundary layer eddies (e.g., rolls or cells, Fig. 9-16). Such echoes are common during the daytime over land when the air temperature is above about 10°C and are associated with insects. The insects do not fill the CBL because they tend to fly downward in updrafts (Geerts and Miao 2005). Further information about rolls can be found in the review paper by Young et al. (2002).

Observed rolls are roughly parallel to the mean CBL wind, and left of the geostrophic wind, as predicted for Ekman-layer rolls by Faller and Kaylor (1966), Brown (1970), and others. Deardorff (1972a) found roll-like structures in his LES for weak instability ($-z_i/L = 1.5$) but not for more unstable values, as did Moeng and Sullivan (1994) for similar values. Subsequent observational work, most recently by Weckwerth et al. (1997) suggest rolls that look like lines of cells can exist up to $-2z_i/L = 20$, but not if $Q_H < 30-50$ W m$^{-2}$, with cells occurring in weak winds. Recently, Salesky et al. (2017)
used LES to reveal a transition of large eddies from rolls to cells as $-z_i/L$ increases from zero to $\approx 15–20$. The horizontal scale for rolls and cells over land tends to be $2–3z_i$; since rolls are aligned nearly parallel to the mean CBL wind, and cells occur in weak winds, it can take up to 30 min to an hour for either one to pass by a fixed point. In LES, rolls are affected by periodic horizontal boundary conditions, which constrain their orientation; this problem can be dealt with by using large horizontal domains (e.g., Glendening 1996). Similarly, the choice of subfilter models can have a strong impact (Ludwig et al. 2009). While Faller, Kaylor, and Brown attribute roll formation to inflectional instability, the wind profile is mostly an organizing factor: observed and LES rolls derive most of their energy from buoyancy. Moreover, shear-parallel rolls result in stability studies of constant unidirectional shear with unstable thermal stratification (e.g., Asai 1970), since shear-perpendicular rolls tend to feed energy into the mean flow.

5. Stable boundary layer

a. Introduction

While it has long been known that stable boundary layers are commonly generated during clear-sky nocturnal
conditions or with advection of warmer air over a colder surface, progress in understanding the stable boundary layer (SBL) has lagged that of the CBL because its smaller-scale and weaker turbulence, greater sensitivity to terrain, and more heterogeneous structure have made the SBL more difficult to observe (especially at night) and to simulate.

However, understanding of the SBL and particularly the nocturnal boundary layer (NBL) has increased substantially in the past two decades, especially through observational studies [e.g., CASES-99 (Poulos et al. 2002) and Surface Heat Budget of the Arctic Ocean (SHEBA; Persson et al. 2002; Grachev et al. 2005)]. Thanks to more powerful computers and new subgrid parameterization schemes (e.g., Mason and Thomson 1992; Sullivan et al. 1994; Bou-Zeid et al. 2005 and references therein; Chow et al. 2005), LES have reproduced some features of NBLs with weak or moderate stability (e.g., Kosović and Curry 2000; Beare et al. 2006). Newer subgrid schemes (e.g., Zhou and Chow 2011) and finer grid resolution show promise in preventing the turbulence collapse that has led to failures in LES in more stable conditions (Jiménez and Cuxart 2005). And Sullivan et al. (2016) simulated SBL structures for the Beare et al. (2006) case with \( L \) as low as 25.5 m, with careful attention to the lower boundary and using grid spacing down to 0.39 m. At the same time, observations have revealed important complications not previously recognized and noted below.

Recognizing that fluxes can vary rapidly with height in the boundary layer, Nieuwstadt (1984) reformulated MOST in terms of local fluxes, referred to as local similarity theory, where the characteristic scales are calculated locally instead of based on surface fluxes. Local similarity theory applies to the NBL above as well as within the surface layer. Local similarity theory has been evaluated by Derbyshire (1990), Högström (1996), Basu et al. (2006), Steeneveld et al. (2006), and others. Grachev et al. (2013) verified local similarity in the surface layer for stable boundary layers after removing cases where the turbulence did not include an inertial subrange.

The turbulence in the NBL has been often classified as either weakly stable or very stable (Mahrt 2014, and citations therein). Van de Wiel et al. (2003) successfully partitioned the NBL into a turbulent regime (roughly weakly stable), an intermittent regime, and a radiative regime where the turbulence remained extremely weak (a subset of the very stable class). A number of modeling studies have found transitions between weakly stable and very stable boundary layers to be dynamically unpredictable (e.g., McNider et al. 1995; Derbyshire 1999a), which might be partly related to thermal interaction with the ground surface and vegetation (van de Wiel et al. 2002). We organize our thinking by contrasting the weakly stable and very stable boundary layers, with the understanding that such a classification greatly oversimplifies the physics of SBLs.
b. Weakly stable boundary layers

The weakly stable NBL is well defined, with turbulence decreasing with height to near zero, often at a local wind maximum, which may also correspond to the top of the nocturnal inversion (e.g., Blackadar 1957), all of which have been used to identify the NBL top. The weakly stable boundary layer is associated with significant winds and/or cloudy skies, and can reach hundreds of meters in depth.

The weakly stable NBL can be vertically partitioned into the roughness sublayer, the surface layer and the outer layer, the first two of which have been discussed in section 3. The outer layer extends from the surface layer top to the NBL top. Boundary layers that are relatively stationary for extended periods can occur during the polar night and be described by Ekman dynamics in the outer layer (Grachev et al. 2005). Turbulence in the outer layer may also follow Rossby similarity theory (Sorbian 1989). Nonetheless, the role of the Coriolis parameter in the nocturnal boundary layer has been generally difficult to establish from observations.

With weakly stable conditions, the interior of the outer layer may be nearly well mixed, but the surface layer and the top of the boundary layer remain stably stratified. Turbulence may be driven partly by a nocturnal jet. Such jets are sometimes driven by inertial effects (Blackadar 1957), baroclinity associated with cooling over sloped terrain (Holton 1967), or a combination of the two (Du and Rotunno 2014; Shapiro et al. 2016). Lundquist (2003), Banta et al. (2006), Cuxart (2008), van de Wiel et al. (2010), Fedorovich et al. (2017), and others provide additional insight into the evolution of the nocturnal jet. The turbulence can become weak or nearly vanish at the nose of the jet where the speed shear vanishes. Turbulence may also be generated by shear above the jet (Conanga and Cuxart 2006).

With winds sufficiently strong, and terrain sufficiently gentle, airflow follows the surface without changing potential temperature, implying continued coupling between the surface and atmosphere (LeMone et al. 2003). For slightly weaker winds, airflow is coupled to the higher terrain features but more detached from lower-elevation sites (Acevedo and Fitzjarrald 2001; Fiebrich and Crawford 2001).

c. Very stable boundary layers

The very stable NBL results from weak large-scale flow and clear skies and remains poorly understood. The boundary layer top can be difficult to define and the zone of strongest generation of turbulence can be semi-detached from the surface or occur in layers (Balsley et al. 2003). Very stable boundary layers are generally quite thin such that the fluxes may vary significantly between the surface and typical observation levels.

The vertical divergence of the radiative flux (clear air) can become important (Edwards 2009), and fog formation is common at some geographical locations. Typical nonstationarity of the very stable regime prevents physical description by similarity theory (Derbyshire 1999b; Mahrt 1999). Such nonstationarity is partly due to submeso28 motions on scales just larger than the turbulence and includes wavelike motions, microfronts, quasi-horizontal meandering motions (Mortarini et al. 2016), and more complex modes. Examples are shown in Fig. 9-17. Because the wind direction cannot be predicted even on short time scales, dispersion models can be of limited skill in the very stable boundary layer. Unfortunately, these are the situations where the concentration of contaminants may be greatest.

The scale of submeso eddies can actually overlap with the largest turbulent eddies, compromising efforts to separate the turbulence from submeso motions in time series. Submeso motions nominally extend up to horizontal length scales of 2 km (the smallest mesoscale motion) although this partitioning is not based on any physics. A number of recent studies have made progress in understanding the different types of turbulent eddies and their relationship to submeso motions (Vercauteren and Klein 2015; Kang et al. 2015; Vercauteren et al. 2016), although a unified theory has yet to emerge.

The intermittency of the turbulence in very stable conditions is generally related to submeso motions or local shear instability. With increasing numerical resources, shear instability and resulting intermittency has become a growing area of research activity (e.g., Flores and Riley 2011; Ansorge and Mellado 2014; Sullivan et al. 2016).

Even shallow slopes or weak surface heterogeneity, of little consequence with more significant winds, can strongly affect the flow with large stability when the turbulent eddies are weak and of particularly small scale. In this regard, most of Earth’s land surface becomes heterogeneous in strongly stable conditions. Thin drainage flows down local slopes can lead to a wind maximum only a few meters above the surface, such that the vertical gradients of wind speed and momentum flux can reverse sign near the surface. The vertical profile of the turbulence intensity may have a minimum at the wind maximum even close to the surface (Horst and Doran 1988; Grisogono et al. 2007; Mahrt et al. 2014).

28 The term may have first been used in an atmospheric science context by Mestayer and Anquetin (1995), but it was later defined to be between turbulence and smallest mesoscales (e.g., see review of the SBL by Mahrt 2014).
Such thin drainage flows are characterized by relatively large stress divergence below their wind maxima (Nadeau et al. 2013).

Most of Earth’s surface is characterized by surface heterogeneity that occurs simultaneously on different horizontal scales. Horizontal temperature advection may be locally important with low wind speeds (Cuxart et al. 2016). Cold pools can form even in gentle terrain, and commonly form in valleys with shallow down-valley slope. The resulting horizontal temperature variance in larger watersheds (airsheds) can increase through the night to values much larger than those for the weakly stable ABL over flat terrain (Fig. 9-18). In such weak winds, the long-known dependence of air temperature on surface cover (e.g., Cornford 1938) would be expected to be stronger. The time required after sunset to reach maximum horizontal variation of observed 2-m temperature at well-exposed sites is a function of array width, horizontal variation of even mild terrain, and station siting. For example, the timing of maximum variability at well-exposed sites in Fig. 9-18 (top) is linked to the travel time of horizontal drainage winds from the edges of the Walnut River watershed. In contrast, low-lying sites sheltered from the wind are influenced more by radiative cooling. Local (of scale < 1 km) cold pools with shallow slopes are often intermittent (Geiss and Mahrt 2015). Cold pools are better defined and thus better understood in uncommon completely enclosed depressions (Whiteman et al. 2004).

d. Representing the SBL in models

Models accommodate uniform parameterizations (such as existing similarity theory) better than class-dependent parameterizations. However, similarity theory suffers from lack of understanding of the underlying basic physics for the SBL. Specification of the turbulent Prandtl number $Pr_T$ as a function of stability as defined by the gradient Richardson number,

$$\text{Ri}_g = \frac{g}{\nu} \left( \frac{\partial U}{\partial z} \right)^2 + \left( \frac{\partial V}{\partial z} \right)^2 - 1,$$

is one example of this difficulty. Some insight is provided by the equation for total turbulent energy that combines.
turbulent kinetic energy and turbulent potential energy
\[
\frac{1}{2} \left( \frac{\partial \theta}{\partial z} \right)^2, \text{ where the Brunt–Väisälä frequency}
\]
\[
N = \left[ \frac{g}{\rho_0} \left( \frac{\partial \theta}{\partial z} \right) \right]^{1/2}, \text{ Mauritsen et al. 2007;}
\]
\[
Zilitinkevich et al. 2013. \right]
\]
This approach is no longer constrained by a critical Richardson number for differentiating between turbulence and laminar flow but rather leads to a transition Richardson number that divides the weakly stable and strongly stable regimes.

Some higher-order ABL models (e.g., Mauritsen et al. 2007; Zilitinkevich et al. 2013; Sukoriansky and Galperin 2013) predict that \( Pr_T \) increases with \( Ri_e \), which has some observational support (e.g., Strang and Fernando 2001; Yague et al. 2006) and has been attributed to inclusion of internal gravity waves. However, there is considerable uncertainty, since gravity wave scales present multiple sampling problems (see Acevedo and Mahrt 2010 and references therein). In addition, the time series near the surface show a complex mix of motions, and clean monochromatic waves are uncommon (Sun et al. 2015). Although self-correlation predicts that \( Pr_T \) increases with \( Ri_e \), Anderson (2009) developed an approach that mitigates the effects of self-correlation and still found that \( Pr_T \) increases with \( Ri_e \).

As noted by Grachev et al. (2007), the influence of other factors such as radiative flux divergence needs to be investigated. Alternatively, modeling approaches, such as that of Li et al. (2015), which emphasizes the physics behind the relationship between \( Pr_T \) and \( Ri_e \), including the different relaxation times for the heat and momentum fluxes and their dependence on scale, are needed.

The turbulence in the weakly stable boundary layer is generally considered to be more or less continuous and to satisfy similarity theory, but this has become increasingly challenged. Sun et al. (2016), for example, used CASES-99 data to show that fluxes are poorly predicted using MOST for their Regime 2 (Fig. 9-19), which roughly corresponds to the weakly stable NBL, where fluxes are determined by large eddies sensing gradients over their depth (Sun et al. 2012). This regime corresponds to the strong-wind part of the “hockey stick” dependence of \( u* \) or \( V_{TKE} \) on the wind speed, which has been found in numerous other studies. For a few examples, see Andreas et al. (2012) over a sea surface, Vignon et al. (2017) over the Antarctic plateau, and Acevedo et al. (2016) over complex terrain.

6. **The diurnal cycle**

Many studies of the CBL and NBL have been limited to near–steady state, or at least slow evolution, by the
requirement for sufficient statistics to document the contributions of turbulence. As research topics, transitions between states have received attention only relatively recently, despite their recognized importance to many applications. While near steady-state CBLs can exist over the ocean, and weakly stable boundary layers can reach near steady state over the Arctic, the CBLs and SBLs described above are in reality special or end states within the diurnal cycle over land. Transitions take several hours of the diurnal cycle even in rather simple situations. Figure 9-20 indicates, for example, that fully convective low-shear boundary layers typically occur within the 2–4 h after local noon in the Midwestern United States.

Over land, the rising and setting of the sun drives a diurnal cycle of temperature and boundary layer depth. The amplitude of the cycle depends on season, latitude, surface type, land use and soil moisture; it interacts with clouds. Geiger (1966) has examples of documented diurnal $\overline{\varphi}$ variation prior to the twentieth century and discoveries of greater frequency of calm winds at night in the early twentieth century. The diurnal variation of $\overline{\varphi}$ has been less studied, but Geiger (1966, p. 107) shows a graph from Frankenberger (1955) of the “well known” $\overline{\varphi}$ diurnal variation, with peaks during the early morning and evening resulting from evapotranspiration into a shallow boundary layer.

In the classical picture (Ball 1960; Lilly 1968; Garratt 1990), the growth of the boundary layer in the morning is determined by competition between surface buoyancy flux, entrainment of air into the growing boundary layer, and subsidence. From section 4b(2), the entrainment ratio $A_B$ was often assumed to be constant, neglecting shear effects. However, Angevine et al. (2001) found that after sunrise, the air below 50 m warmed to form a nascent boundary layer before $B_0$ became positive, with the CBL growing much faster than suggested for buoyancy-driven growth with $A_B = 0$, indicating the need to account for shear-driven mixing and its role in entrainment. During this “morning transition,” the boundary layer is in a “mixed CBL–SBL state” both vertically, as revealed in LES by Beare (2008) and—in even mildly varying terrain—horizontally, as shown in observations by Lenschow et al. (1979). Only by mid-to-late morning does the buoyancy-driven CBL achieve the entrainment ratios produced by LES of buoyancy-driven CBLs. While direct heating of the air by radiation was recognized as being a factor in daytime warming, its effect over land is often fairly small (e.g., Deardorff 1974a; Angevine et al. 1998).

During the afternoon and evening, the boundary layer was found to undergo “transition” and “collapse” of turbulence by early researchers. The associated reduction in near-surface thermal instability, followed by onset of a nocturnal inversion resulting from radiative cooling on clear nights, has been documented for a century or more (Geiger 1966). Now we have a richer picture of coevolution of turbulence and mean profiles during late afternoon and evening. As the sun moves lower in the sky through the afternoon, the surface cools more rapidly than the air a few meters above the surface, making the air less unstable. Therefore, thermals weaken. They rise through an approximately neutral layer, so even these weaker thermals can reach the same height as earlier; but eventually other processes become important, and most of these processes tend to warm and stabilize the upper BL (e.g., Grant 1997). For example, shear-driven entrainment and direct radiative heating both warm the upper BL. Subsidence also drives the BL top downward. The result is that the CBL depth as defined by an increase in $\partial\overline{w}/\partial z$ stays roughly constant or decreases in the afternoon, with the $w'w''$-defined depth (top of radar-sensed plumes, since $\overline{Ri}$ higher up indicates little local mixing) decreasing even faster on many occasions. Grimsdell and Angevine (2002) were the first to describe an afternoon transition in these terms, drawing from in situ flux profiles (e.g., Grant 1997). Figure 9-21 shows an example of this evolution. The afternoon weakening of turbulence and decline of the BL top can occur slowly enough that statistical stationarity still applies.

---

**Fig. 9-20.** Average diurnal cycle of vertical wind shear [kt (100 ft)$^{-1}$, which corresponds to 1.7 m s$^{-1}$ (100 m)$^{-1}$], from one year of data from the KTVY-TV tower in Oklahoma City, which was instrumented by the National Severe Storms Laboratory. [From Crawford and Hudson (1970).]
Eventually, $\tilde{\theta}_0$ falls below $\tilde{\theta}_{0,\text{air}}$, making the surface layer stable. As a result, $w'\tilde{\theta}_y$ becomes negative and the winds decelerate as downward momentum mixing decreases; these changes coincide with the onset of the evening transition. Businger (1973) argued that this cooling can make the air stable enough that turbulence is suppressed—the so-called collapse. This situation tends to be temporary, since turbulence persists at higher altitudes, resulting in an increase in near-surface wind and thus a resumption of near-surface turbulence once the vertical shear becomes sufficient for turbulence to be re-established. As shown by van de Wiel et al. (2002, 2003), this process can repeat through the night.

The Boundary Layer Late Afternoon and Sunset Turbulence project (BLLAST; Lothon et al. 2014) was dedicated to deeper understanding of these processes. On the modeling side, the Global Energy and Water Cycle Experiment (GEWEX) Atmospheric Boundary Layer Study (GABLS) GABLS2 and GABLS3 single-column modeling comparisons included full diurnal cycles (Bosveld et al. 2014; Svensson et al. 2011). In both studies, the various models produced quite different responses, especially in the evening. Recently, Basu et al. (2008), Kumar et al. (2006, 2010), and Sharma et al. (2017) have conducted LES of the diurnal cycle, providing reasonable matches to observations. Such studies are rare, mainly because of the need for finer grids and sophisticated subfilter-scale representations needed for stable stratification.

### 7. Impact of horizontal heterogeneity

Any time airflow encounters an abrupt change in surface characteristics, the boundary layer will be out of equilibrium and a circulation will develop. The sea breeze is the best-known example, but horizontal heterogeneities also matter over land. From the early twentieth century, researchers have documented the influence of vegetation or soil properties on near-surface temperature, humidity, and wind (Geiger 1966). Starting in the late 1950s, researchers began to examine internal boundary layers resulting from a change in surface roughness; first, for the simple case of neutrally stratified flow, but later examining the effects of abrupt changes in surface flux and temperature (Garratt 1990). Such information was helpful in figuring out just where to measure a surface flux representative of, say, a large cornfield; but for flux measurements from a moving platform like an aircraft, the problem is turned around: just which surface contributed to the measured fluxes?

As can be seen from Fig. 9-22, there is a qualitative relationship of fluxes to surface properties beneath and upstream of the aircraft track, but more quantitative information is desirable for testing surface-flux models. The contribution of the upstream surface to the fluxes measured at a point is determined in terms of a so-called footprint, which is a function of height, stratification, and wind speed. Schuepp et al. (1990, 1992) refined a “footprint” function, derived for diffusion applications by Pasquill and Smith (1983) to address this question, and applied it to aircraft data from FIFE (Sellers et al. 1992). Davis et al. (2003) used the Weil and Horst (1992) footprint formulation to show that the fetch is on the order of 10 times the measurement height during convective conditions, with this increasing significantly during stable conditions. Thus, for a surface-level measurement, a 100–200-m fetch is recommended. If one can model the surface fluxes for each surface type contributing to the measured flux using its flux footprint (e.g., Anderson et al. 2004), then the model can then be applied to similar parcels of land to “scale up” the fluxes over a region (e.g., Anderson et al. 2007), sometimes referred to as the tile approach.

Fluxes over a heterogeneous area can also be computed applying MOST to roughness lengths that apply to the area (Wieringa 1976). These regional or effective roughness lengths apply above a so-called blending height—where the fluxes from the nonuniform surface become blended—but still within the surface layer; that is, the blending height must be low (In the case of Anderson et al. 2004, about 50 m). Estimates of the effective roughness length $z_{eff}$ were perhaps first made when Fiedler and Panofsky (1972) applied a simple...
equation relating the ageostrophic wind to boundary layer fluxes using aircraft wind and turbulence data collected at 75 and 225 m AGL. Ironically, this was the same year that Pasquill (1972) introduced the blending-height concept. Since weather and climate models compute fluxes over grid areas that typically encompass a nonuniform surface, \( z_{\text{eff}} \) and its heat-flux and moisture-flux counterparts \( z_{\text{eff}}^{h} \) and \( z_{\text{eff}}^{q} \) are needed on a routine basis (Beljaars and Holtslag 1991). The blending height has been estimated as proportional to the scale of the heterogeneity (Mason 1988), and varies with stability, wind speed, and surface properties (e.g., Wood and Mason 1991; Mahrt 1996).

Beljaars (1988, see Beljaars and Bosveld 1997) estimated \( z_{\text{eff}}^{h} \) as a function of fetch and season for the 200-m Cabauw tower in the Netherlands and found it to be larger than that at the tower itself; this work was refined using an expanded dataset by Verkaik and Holtslag (2007). Kustas et al. (2005) obtained \( z_{\text{eff}}^{h} \) for the CBL over an area with corn and soybean fields both by using aircraft data collected above the blending height and by aggregating the roughness lengths of the fields in the aircraft flux footprint following Mason (1988) and found that the aggregated value—and therefore the associated flux—was too small. However, they successfully managed to replicate heat and moisture fluxes. They concluded that the too-small \( z_{\text{eff}}^{h} \) was related to flow over sharp boundaries between fields, such as created by shelterbelts, edges of roadways, or the height differences between the two crops, as well as upstream buildings and other obstacles not accounted for. A number of years earlier, an LES study of a neutrally stratified BL by Albertson and Parlange (1999b) revealed that the edges of patches with different roughness contributed to the total surface stress, with stress larger for more, smaller patches. More recently, Bou-Zeid et al. (2004) performed a series of neutral-boundary layer LES for a similar range of patch sizes to develop relationships between patch size, blending height, and \( z_{\text{eff}}^{h} \).

As for \( z_{\text{eff}}^{q} \), Wood and Mason (1991) and Beljaars and Holtslag (1991) both showed that it can be significantly different from local values in the CBL. For a weakly stable BL with continuous turbulence, over a surface with streamwise horizontally varying temperature, Stoll and Porté-Agel (2009) suggest that the scalar fluxes

Fig. 9.22. (a) The influence of surface properties on sensible heat flux \( Q_{h} \), based on aircraft data collected at 65 m AGL, (b) along a track flown during the spring (CASES-97: dormant grass and growing winter wheat) and summer [IHOP_2002, Weckwerth et al. (2004); senescent/harvested winter wheat and green grass]. Fluxes are averages over \( N \) flight legs, of 4-km running-mean averages of instantaneous values of \( Q_{h} \) based on departures of \( \theta \) and \( w \) from their linear trends along each leg. Dashed lines are \( \pm \sigma(N - 1)/N^{0.5} \), where \( \sigma \) is standard deviation. Wind direction is from the south. [Panel (a) and top of (b) are adapted from Figs. 5–7 of LeMone et al. (2007).]
found from averaging MOST-derived fluxes for individual patches (\(~\text{one-half to twice PBL depth}\) grossly underestimate the actual heat flux, unless local similarity theory is used. For further information on blending-height theory and related concepts, see Mahrt (2000).

Relating fluxes to surface properties is made possible by databases for land cover and soil properties, which are heavily dependent on satellite remote sensing. This of course required translating measurements at different wavelengths into vegetation properties, such as the leaf area index (LAI, leaf area per unit surface area, estimated first by Jordan 1969), and the normalized differential vegetation index (Rouse et al. 1974), a measure of vegetation greenness. The first land-use satellite (Landsat) was launched in 1972; the first global land-cover database was completed in the 1990s (https://archive.usgs.gov/archive/sites/landcover.usgs.gov/globallandcover.html; also see Loveland et al. 1995). In the United States, two commonly used land surface datasets are based on Advanced Very High Resolution Radiometer (AVHRR) data, provided by the U.S. Geological Survey, and the Moderate Resolution Imaging Spectroradiometer (MODIS). Soil moisture and temperature are found through use of land surface models (see section 10).

There emerged in the 1980s an increased interest in the impact of surface heterogeneity on boundary layer structure. Horizontal nonuniformity in soil moisture and land cover and even modest changes in elevation lead to mesoscale circulations that are more than curiosities: they can initiate development of thunderstorms (e.g., Pielke et al. 1997; Fig. 9-23). Such an association was anticipated in idealized mesoscale simulations by Ookouchi et al. (1984). Numerous other authors have examined the impact of horizontal nonuniformities using mesoscale models (Yan and Anthes 1988), LES (Walko et al. 1992), and LES and DNS (Krettenauer and Schumann 1992). These circulations vary in horizontal extent from \(~5\) km (Mahrt et al. 1994) to \(~100\) km (Chen and Avissar 1994b). Observations are scarce, although Physick and Tapper (1990) document diurnal circulations around a \(~8\)-km diameter salt pan in Australia, LeMone et al. (2002) find 60-km circulations for a summer day when dormant (warm) vegetation at higher elevations reinforced terrain effects in forcing the circulations, and Mahrt and Ek (1993) found evidence of “forest breezes,” in HAPEX-MOBILHY in France, almost 100 years after they were first reported (Schmauss 1920; see Geiger 1966).

How do these structures impact fluxes and profiles? Numerous LES studies have addressed the impact of surface heterogeneity and associated circulations on regional boundary layer statistics. Comparing a simulated CBL over a uniform surface to one with imposed realistic surface temperature variations of scale 450–900 m to represent the surface over which observations were taken, Hechtel et al. (1990) found little difference between simulated boundary layer evolution and horizontally averaged variances \(\theta, q\), and \(w\), buoyancy flux, \(q\) flux and TKE, which, they noted could partially result from a \(~5\) \(\text{ms}^{-1}\) wind. Avissar and Schmidt (1998) found that idealized but realistic heat-flux patchiness of scale \(~5–10\) km had little effect on the average CBL \(\theta\) and \(q\) statistics, but they did find impact for larger scales and winds \(< 5\) \(\text{ms}^{-1}\). Similarly, Albertson and Parlange (1999a) find the blending-height concept to be useful for imposed heat and moisture fluxes on scales on the order of PBL depth owing to mixing by PBL eddies of comparable scale to \(z_i\), especially for mean profiles of \(\theta\) and \(q\).

Cities are important sources of horizontal heterogeneity throughout the boundary layer. Oke (1982) traced our awareness of urban–rural differences in near-surface temperature back to the early 1800s but notes the absence of systematic investigation going beyond mere statistical association until around 1970. By the 1980s, a number of studies (e.g., Bornstein 1968, Changnon 1981) revealed that, compared to the surrounding rural areas, cities at night tend to be warmer near the surface,
with $\bar{b}$ less stratified; while during the day, the urban CBL tends to be deeper. Oke attributed these changes not only to the impact of different structure and composition of cities on heating, cooling, and momentum exchange, but also to release of heat from buildings, factories, cars, and even people. Fernando (2010) and Barlow (2014) provide more recent reviews of the effects of cities on the ABL.

8. Cloud-topped boundary layers

a. Introduction

Cloud-topped boundary layers (CTBLs), especially over the oceans, have developed into one of the most important topics in boundary layer meteorology because of their extensive coverage and climate impact. Clouds strongly interact with solar and infrared radiation and they can precipitate; consequently, they affect and are affected by the surface energy balance as well as the ABL turbulence structure. They also strongly impact ABL aerosols and chemical composition. Overall, they significantly impact Earth's radiation balance and general circulation. The response of boundary layer clouds to anthropogenic greenhouse gases and aerosols are two of the key modeling uncertainties in projecting climate change (IPCC 2013, chapter 7). CTBLs are ubiquitous: based on routine surface observations compiled in the Warren cloud atlas (Warren et al. 1986, 1988), on average over 47% of the globe (55% of the ocean and 26% of the land), is covered by low-lying cloud cover with a base below 1 km (Hahn and Warren 1999; Fig. 3 in Park and Bretherton 2009), most of which is likely rooted in the ABL.

While interest in the tropical marine CTBL can be traced back to the 1930s, research on CTBLs picked up during the 1980s, as a natural next step after rapid progress on the clear-air CBL (section 4). The timing was good, since NWP and climate models were just beginning to predict cloud development. This effort was supported by a series of international field programs with a focus on cloud-topped boundary layers, taking advantage of the latest developments in remote sensing, surface flux, and aircraft turbulence and microphysics measurement technology. Also, starting in 1992 with the Atmospheric Radiation Measurement (ARM) Program Southern Great Plains (SGP) facility, a series of surface “supersites” in different parts of the world became available to provide valuable surface-based observations to complement such field efforts. The GEWEX project began to foster a series of LES and single-column model (SCM) studies of significant CTBL regimes and processes sampled in field campaigns, some initiated by the climate community. Thanks in large part to this synergy, LES technology improved to the point that it could be used as a benchmark for developing CTBL parameterization schemes for increasingly complex scenarios.

b. Cumulus-topped boundary layers

Kite soundings from the Meteor cruises of the 1930s (von Ficker 1936) established the vertical wind and temperature structure of the Atlantic trade winds. After World War II, airborne meteorological research focused on the relatively benign nonprecipitating trade wind–cumulus boundary layer regimes that were accessible over the island-dotted Caribbean and west Pacific. Stommel (1947) was the first to quantitatively estimate the rate of lateral mixing of air into a shallow cumulus cloud. Bunker et al. (1949), Malkus (1958), and Malkus and Riehl (1964) distilled aircraft observations to document the vertical structure of trade wind–cumulus boundary layers and noted their mesoscale organization into rows and clusters.

As noted in section 4, ATEX, BOMEX, and GATE in the 1960s and 1970s documented profiles and fluxes for several shallow-cumulus CTBLs, which formed the basis for studies using the emerging LES technology. Sommeria (1976) showed realistic cloud profiles and turbulence statistics in a pioneering LES of nonprecipitating shallow cumulus clouds idealized from BOMEX and other contemporary field studies; and did a reasonable job of reproducing subcloud fluxes and structure for a fair-weather trade–cumulus case sampled north of Puerto Rico (Sommeria and LeMone 1978). Indeed, BOMEX measurements for 22–26 June 1969 featured a quasi-steady nonprecipitating trade cumulus boundary layer that was used as a GEWEX Cloud System Study (GCSS) benchmark case for testing models and parameterizations in this regime (Siebesma et al. 2003). As in the case of the Moeng and Sullivan (1994) LES clear-air buoyancy-driven CBL in Fig. 9-10, $A_B \approx 0.2$ in their simulation if one uses cloud base height as $z_i$. In an LES of a small-Cu case with geostrophic wind $\sim 10 \text{ m s}^{-1}$ over Oklahoma, Brown et al. (2002) obtain $A_B \approx 0.15$. This is also true for the GATE cases (Fig. 9-13), for which $A_B = 0.17$ and the GATE TKE budget (Fig. 4 in LeMone 1980) resembles the buoyancy-driven CBL TKE budgets in Fig. 9-12. Thus, despite the presence of cloud “roots” (Fig. 9-15) and the strong association of $\vec{w}' q'$ and $\vec{w}' b'$ profiles with cumulus cloudiness (Fig. 9-13), the CBL TKE budget in the presence of weak nonprecipitating trade cumuli is to a good approximation decoupled from the clouds.

For more intense clouds and/or on larger scale, coupling between clouds and the subcloud layer starts to manifest itself at least in terms of mesoscale structure.
LeMone and Meitin (1984) used GATE aircraft data to document mesoscale patches and bands of concentrated shallow to moderate cumulus 10–30 km across that lay above moister regions in the subcloud layer. Balaji et al. (1993) later associated the cumulus bands for one of the cases with modulation of the CBL by tropospheric gravity waves, launched as “convection waves” over the boundary layer clouds but eventually assuming a wavelength consistent with their stable environment, as described in Clark et al. (1986). Their flatness (band spacing several times ABL depth) is consistent with the cold-air outbreak “longitudinal rolls” at lower latitudes in Fig. 6 of Miura (1986) and the “sea cold-air outbreak” rolls in Fig. 4 of Young et al. (2002). Melfi and Palm (2012) postulate the same mechanism for such cloud bands in a cold-air outbreak over the North Atlantic. These broad cloud bands are not to be confused with the more closely spaced cloud streets observed over the tropical ocean, which lie above the clear-air convective rolls discussed in section 4. However, our knowledge of boundary layers with convective clouds of intermediate development remains the most incomplete, largely because such BLs are so difficult to sample and precipitation becomes more significant.

c. Stratocumulus-topped ABLs

1) ENTRAINMENT AND THE STRATOCUMULUS–CUMULUS TRANSITION

Radiosonde data collected from three weather ships stationed between the U.S. West Coast and Hawaii during World War II provided Riehl et al. (1951) the opportunity to document the downstream development of the northeast Pacific CTBL between the cool waters off California and the warmer waters near Hawaii. They noticed a close connection between the vertical structure of the CTBL and the associated cumulus clouds (Fig. 9-24), and mentioned the stratiform clouds closer to the California coast. Using the along-cross-section wind divergence to estimate the mean vertical motion in this region, they demonstrated that free-tropospheric air must be descending through the inversion and becoming turbulent mixed or entrained into the cloud layer, weakening the inversion toward the west and increasing its height. A similar result was obtained by Neiburger et al. (1961), who used radiosonde data from postwar research cruises to map out the offshore increase in inversion height and decrease in horizontal divergence under summertime conditions. For more detail of the history of marine stratocumulus experiments off the California coast, see Kloesel (1992).

Intrigued by the shallow stratocumulus-topped boundary layers capped by a strong sharp inversion, which is prominent in summertime radiosonde observations along the California coast, Lilly (1968) developed a seminal mixed-layer model, based on the insight that the turbulence that mixes moisture upward to form clouds was driven by infrared-radiation cooling at cloud top. The turbulence also induces entrainment of warm, dry air from above that regulates the cloud thickness and boundary layer depth. Lilly’s model demonstrated for the first time the tight feedbacks among clouds,
radiation, turbulence and surface fluxes characteristic of many CTBLs. His paper used the TKE budget drawing on concepts introduced by Ball (1960) to address the difficult problem of entrainment closure, which has remained a theme of stratocumulus-topped boundary layer research for the subsequent 50 years.

The 1976 Marine Stratocumulus Experiment (Wakefield and Schubert 1976; Brost et al. 1982a,b; Albrecht et al. 1985) utilized an aircraft off the California coast in the first extensive observational test of Lilly’s mixed-layer model. Wakefield and Schubert (1981) applied a slightly generalized form of the Lilly model to the Lagrangian evolution of stratocumulus (Sc) along typical summertime boundary layer trajectories and found that the Sc layer tended to thicken indefinitely downstream. Thus, this model could not explain the Sc–cumulus (Cu) transition. Randall (1980) and Deardorff (1980b) proposed that the Sc–Cu transition was due to cloud-top entrainment instability, a runaway feedback between entrainment-induced evaporative cooling and turbulence production in a Sc cloud layer; radiosonde measurements taken as part of the First ISCCP Radiation Experiment (FIRE; Albrecht et al. 1988) disproved the original criterion (Kuo and Schubert 1988), but more restrictive versions are plausible (e.g., Lock 2009).

Given the important role of entrainment in Sc evolution, observational estimation of the small (typically less than 1 cm s\(^{-1}\)) but important entrainment rate and its relation to other cloud and boundary layer properties became a priority. Nicholls and Turton (1986) and others inferred entrainment rate from airborne measurements using a humidity flux/jump method, and suggested a relationship between entrainment and turbulence within the Sc layer; uncertainties were too large to rule out other possible entrainment relationships, however. Airplane observations from the Dynamics and Chemistry of Marine Stratocumulus (DYCOMS)-I in 1985 (Lenschow et al. 1988) demonstrated the applicability of trace species (e.g., ozone) for measuring entrainment (Kawa and Pearson 1989), and later, dimethyl sulfide (DMS) during DYCOMS-II (Faloona et al. 2005). As described in the review article by Mellado (2017) and more recent papers by, for example, Shi et al. (2018a,b) and van Hooff et al. (2018), advanced DNS and LES should provide further insight into entrainment at the top of stratocumulus clouds.

2) DIURNAL CHANGES

Using summertime aircraft observations over the North Sea, Nicholls (1984) showed that daytime in-cloud solar heating tends to stratify a stratocumulus-topped boundary layer, leading it to “decouple” into separate layers, one surface-driven and the other driven by cloud-top longwave cooling. The two layers can then “recouple” into a single mixed layer overnight.

Since low-level aircraft flights were made mostly during daylight, surface-based remote sensing opened new windows on the diurnal cycle of stratocumulus. Tethered balloon studies of nocturnal stratocumulus over the United Kingdom elucidated the strong vertical gradients in turbulence, cloud properties, and thermodynamic structure near the capping inversion (Roach et al. 1982; Caughey et al. 1982). During FIRE in June 1987, a laser ceilometer, a microwave radiometer, and an acoustic sounder on an island off the California coast documented the diurnal cycle of cloud, boundary layer depth, and daytime boundary layer decoupling (Albrecht et al. 1990). During the June 1992, Atlantic Stratocumulus Transition Experiment (ASTEX; Albrecht et al. 1995), millimeter-wavelength cloud radars on two islands revealed surprising prevalence of drizzle from stratocumulus, especially at night (Miller and Albrecht 1995; Frisch et al. 1995).

3) DECOUPLING AND THE STRATOCUMULUS–CUMULUS TRANSITION

In the early 1990s, an alternative to the Wakefield–Schubert Sc–Cu transition hypothesis was developed, inspired by Nicholls’s work on diurnal decoupling. It was based on turbulence closure modeling (Bretherton 1991) and mixed layer modeling (Bretherton and Wyant 1997) of the Lagrangian evolution of the Sc layer as it moved over warmer waters, and later, two-dimensional LES-like simulations of this multiday transition (Krueger et al. 1995a,b; Wyant et al. 1997). As a Sc-capped mixed layer moves over warmer waters, the capping inversion weakens, leading to more entrainment and a rising capping inversion. The increased entrainment ultimately causes the Sc mixed layer to decouple from the surface. Cumulus clouds form on top of the resulting surface mixed layer, and rise up into the overlying Sc mixed layer. As the inversion rises, the cumulus clouds become deeper and more vigorous, and their overshooting tops entrain increasingly more air across the inversion. This disrupts the stratocumulus layer, ultimately leaving behind only a trade cumulus layer over warm waters (Wyant et al. 1997). Another contributing factor to the stratocumulus dissipation may be the humidity jump between the stratocumulus and the dry overlying air, which becomes increasingly large over warmer water and may contribute to an evaporative enhancement of the efficiency of entrainment (Lock 2009). ASTEX pioneered multiday airborne Lagrangian sampling of the Sc–Cu transition (Bretherton et al. 1995a,b), creating unique datasets that
largely confirmed this new hypothesis (Fig. 9-25), and which are still invaluable for process and LES intercomparisons (Bretherton et al. 1999a; van der Dussen et al. 2013).

4) CLOUDS, AEROSOLS, AND PRECIPITATION

Early field studies of marine boundary layer clouds frequently encountered drizzle and rain showers even from clouds only a few hundred meters thick (e.g., DYCOMS-I; Lenschow et al. 1988), and sampled heavier precipitation on the windward side of islands in the trades like Hawaii [e.g., the 1990 Hawaiian Rainband Project (HaRP); Paluch et al. 1994]. Likewise, precipitation was a common feature of trade wind CTBLs in ASTEX (Albrecht et al. 1995).

In the first airborne deployment of a Doppler cloud radar and lidar, Vali et al. (1998) found extensive precipitation in a marine Sc layer less than 300 m thick off the Oregon coast. Ship-based cloud radar over the southeast Pacific stratocumulus region in the East Pacific Investigation of Climate (EPIC; Bretherton et al. 2004b) and airborne cloud-radar observations during the June–July 2001 DYCOMS-II (Stevens et al. 2003), together with satellite-based estimates of droplet size, suggested that regions of broken Sc within solid Sc decks, called “pockets of open cells” (POCs) or “rifts,” were regions of more cumuliform and heavily precipitating cloud with lower aerosol concentrations (Fig. 9-26; Stevens et al. 2005b). Penetrations of numerous POCs in the VAMOS Ocean Cloud Atmosphere Land Study (VOCALS) Regional Experiment (2008; Wood et al. 2011b) confirmed this view and documented an “ultraclean layer” of highly depleted aerosol and very low cloud droplet concentrations just below the trade inversion base, maintained by precipitation scavenging of aerosol (Wood et al. 2011a;
LES in mesoscale domains showed the importance of cloud–aerosol–precipitation–dynamics interaction in maintaining long-lived POCs embedded in solid stratocumulus (Wang et al. 2010; Kazil et al. 2011; Berner et al. 2013). In January–February 2005, the Rain in Cumulus over the Ocean experiment (RICO; Rauber et al. 2007), focused on precipitation formation in tropical Atlantic shallow cumulus clouds and its interplay with cloud organization, boundary layer depth, and aerosols. GEWEX Cloud System Study (GCSS) LES intercomparison studies of both precipitating stratocumulus (Ackerman et al. 2009) and shallow cumulus (van Zanten et al. 2010) showed considerable sensitivity of simulated precipitation and cloud properties to the choice of microphysics parameterizations; this remains a key uncertainty in LES and SCM of precipitating boundary layer clouds and their response to aerosol perturbations. LES suggests that precipitation may often control the depth of trade cumulus boundary layers (Blossey et al. 2013; Seifert et al. 2015).

5) COLD-AIR OUTBREAKS

AMTEX (Lenschow and Agee 1976) documented airmass modification and cloud structure in midlatitude oceanic cold-air outbreaks, another key marine boundary layer regime. Using ships, buoys, aircraft and carefully collocated satellite imagery, open and closed mesoscale cellular cloud organization (Agee et al. 1973) were sampled and related to the associated boundary layer profiles and turbulent moisture and buoyancy fluxes (Burt and Agee 1977; Rothermel and Agee 1980). As was the case for the trade–cumulus BL, The AMTEX TKE budget (Lenschow et al. 1980) resembled the Moeng and Sullivan (1994) budget for a buoyancy-driven CBL (Fig. 9-12), despite the presence of
scattered-to-broken stratocumulus clouds and strong baroclinicity.

Recent GCMs produce systematic cloud biases around midlatitude oceanic cyclones, with too little cloud in their cold sectors, which are dominated by boundary layer stratus cloud transitioning into shallow cumuli (Bodas-Salcedo et al. 2014). A Global Atmospheric System Study (GASS) intercomparison based on a North Atlantic cold-air outbreak (Field et al. 2017) showed that LES and 1-km grid regional models also have these biases. One issue may be premature glaciation and rapid dissipation of supercooled liquid stratus cloud, possibly in association with excessive ice nucleation and rapid dissipation of supercooled liquid stratus cloud, possibly in association with excessive ice nucleation. Parameterization changes inhibiting cloud glaciation at higher temperatures have been developed to control this bias; boundary layer observations in this regime were taken in the January–February 2018 Southern Ocean Cloud Radiation Aerosol Transport Experimental Study (SOCRATES) south of Tasmania to test if other microphysical, aerosol, or boundary layer processes are also important.

d. Improved representation of marine boundary layer clouds in global models

GCSS was started in 1992 (GEWEX Cloud System Science Team 1993; Randall et al. 2003) in order to more effectively use observations to improve large-scale and process models. The GCSS Boundary Layer Cloud working group designed a series of intercomparisons for LES and SCM versions of weather and climate prediction models, distilled from observational case studies. After a first nocturnal stratocumulus case with discouragingly divergent results between models (Moeng et al. 1996), cases were chosen to isolate important modeling issues, initially numerical algorithms and grid resolution, while using idealized formulations of other physical processes such as radiative cooling that could be standardized across models. A “smoke cloud” case (Bretherton et al. 1999b) inspired by Lilly (1968) was even compared with a laboratory experiment rather than field observations.

By the early 2000s, LES representation of observed CTBLs was getting good enough that they could realize their potential as benchmarks for informing parameterization development. GCSS successfully tackled a shallow cumulus development over land (Brown et al. 2002) idealized from a June 1997 case over the ARM SGP site. In both this and the earlier BOMEX case, different LES run at the same grid resolution gave generally similar results. However, GCSS intercomparisons showed that more sophisticated advection schemes and 5–10-m grids were necessary to avoid overentrainment and premature breakup of Sc clouds, not a surprising result since DYCOMS-I observations by Lenschow et al. (2000) indicated that the discontinuity in thermodynamic variables at cloud top could be less than 1 m thick. With a 5-m vertical grid near the inversion, a subset of LES reasonably matched the observed Sc thickness, entrainment rate, and turbulence characteristics for a nonprecipitating Sc-capped mixed-layer case study in DYCOMS-II (Stevens et al. 2005a; Zhu et al. 2005). More recent and comprehensive studies have shown that LES with 5-m vertical grid, interactive radiation, microphysics and surface flux parameterizations, and accurate large-scale forcing can also skillfully represent observed cloud and boundary layer characteristics in the subtropical Sc–Cu transitions (van der Dussen et al. 2013; McGibbon and Bretherton 2017).

In the GCSS stratocumulus and shallow cumulus studies, SCMs showed a broad range of behaviors. This, together with the large marine boundary layer cloudiness biases in global and regional models, motivated intensive work on marine boundary layer parameterization development, leading to improved simulations of stratocumulus and shallow cumulus regimes in weather and climate models (e.g., Lock et al. 2000; Golaz et al. 2002; Bretherton et al. 2004a; Siebesma et al. 2007; Bretherton and Park 2009; Park and Bretherton 2009; Neggers et al. 2009; Köhler et al. 2011).

9. Boundary layer over the Arctic Ocean

While the physics governing the Arctic atmospheric boundary layer (AABL) is fundamentally the same as everywhere else, the AABL is set apart from ABLs at lower latitudes by a unique combination of forcings as a consequence of its high latitudes, and a strong influence of snow and ice at the surface. The inaccessibility and harsh conditions of this remote region combine to make field work in the Arctic difficult. Consequently, fewer AABL observations are available than for ABLs at lower latitudes.

a. A brief look back in time

The history of AABL research goes back to the late 1800s and early 1900s, when the early explorers were interested in the atmosphere because of its influence on the polar ocean and in particular the sea ice. While many of these expeditions, like Fritjof Nansen’s Fram expedition, only took standard surface observations (Mohn 1905), these were of sufficiently high quality to be used today. Some early expeditions, such as Amundsen’s 1918–21 expedition on the Maud, also took vertical profiles using

---

29 A problem also in modeling of Arctic stratocumulus (e.g., Prenni et al. 2007; Pithan et al. 2014).
b. Developing a scientific understanding

Two principal factors contribute to shaping the AABL. First—and perhaps most obvious—is the relatively smaller role of solar radiation in the local surface

Laboratory balloons or kites. As described by Sverdrup (1922), Amundsen’s aim was to drift across the Arctic Ocean, like Nansen on the Fram. While forced to spend three winters in the Arctic because of difficult ice conditions, they took observations during the whole expedition, including vertical profiles that revealed the strong surface temperature inversions that sometimes occur in winter (Sverdrup 1922, 1925). All the data from the expedition were delivered intact once the Maud returned. However, an attempt to safeguard observations after the first winter by sending copies overland led to the deaths of the two men carrying the data (Sverdrup 1922). This illustrates both the dedication of these early explorers and the harsh and unpredictable conditions they encountered.

After the early explorers, the literature reveals little activity during the first half of the twentieth century. The expeditions that were deployed usually had little or no atmospheric component. Although not being an AABL project per se, the manned Russian Drifting Stations, also called Ice-islands or “North Pole” (NP) stations (Romanov et al. 1997), deserve attention. These stations were launched annually between 1937 and 1991; the program was resurrected in 2004 but ended again in 2015. The NP stations performed regular soundings that today form a backbone for our knowledge about the vertical structure of the AABL (e.g., Kahl et al. 1996, 1999).

Even as late as the 1970s the primary goal for micro-meteorological observations during the Arctic Ice Dynamics Joint Experiment (AIDJEX; Untersteiner 1979) was to determine how momentum flux from the atmosphere caused sea ice to move around and deform (Maykut et al. 1972; Arya 1979). While AIDJEX was primarily a sea ice experiment, several AABL observations were taken, including surface-layer temperature profiles, wind speed profiles, and eddy-correlation fluxes, mostly to derive formulations for $z_{0m}$ and $C_D$ for Arctic sea ice. AIDJEX was also one of the first experiments deploying research aircraft in the Arctic and one of the first focusing on AABL clouds. Several smaller aircraft-based observation campaigns followed, for example the Arctic Stratus Experiment in 1980 (e.g., Curry and Herman 1985) and the Lead Experiment (LEADEX; Morison et al. 1993) in the early 1990s, marking a shift in emphasis from sea ice to the AABL itself. For example, LEADEX studied the spectacular turbulent fluxes over open water leads in winter. Several aircraft-supported experiments studied the AABL transition in the marginal ice zone (MIZ) in on- and off-ice flow, including the 1983–84 Marginal Ice Zone experiment in 1984 (Morison et al. 1987; Kellner et al. 1987), the 1991 and 1993 Radiation and Eddy Flux Experiments (REFLEX; Hartmann et al. 1992; Kottmeier et al. 1994), and the 1997–2000 Arctic Radiation and Turbulent Interaction Study (ARTIST; Hartmann et al. 1999).

An exciting new observing platform was introduced in 1980 with an international expedition based on the Swedish icebreaker Ymer, which provided a safe platform for staff as well as expensive, sensitive, and heavy instruments that did not need to be deployed on the ice. Organized in commemoration of Nordenskjold’s voyage through the Northeast Passage (e.g., Schytt 1983), the expedition’s measurements included atmospheric chemistry and aerosols, with AABL observations made both onboard, and by a shore team on Spitsbergen. Shorter-term icebreaker-based research is typically conducted in summer or autumn, when the ice is easier to navigate. Since the early 1990s, several such AABL campaigns were undertaken with the Swedish icebreaker Oden (e.g., Leck et al. 1996, 2001; Tjernström et al. 2004, 2014; Sotiropoulou et al. 2016) and the German icebreaker Polarstern (e.g., Driemel et al. 2016; Lüpkes et al. 2010). The Canadian icebreaker Amundsen has also been involved, for example, the 2007–08 Circumpolar Flaw Lead Study (Barber et al. 2010), however, with only a modest AABL component.

Some more recent Arctic field campaigns have had a strong focus on the AABL. Perhaps the most significant to date is the 1997–98 SHEBA campaign (Uttal et al. 2002; Persson et al. 2002), when the Canadian icebreaker Des Groseilliers was frozen into the ice north of Alaska and drifted with it for a full year. While SHEBA was multidisciplinary, its substantial AABL program has become a gold standard for AABL studies, with eddy-correlation flux measurements, mean profiles, soundings, and remote sensing observations of AABL clouds. Indeed, SHEBA generated an extensive record of stably stratified turbulence, providing insight for similar conditions also outside of the Arctic (e.g., Grachev et al. 2005; Mauritsen and Svensson 2007; Holtslag et al. 2013, also see section 5). During the 2007–08 International Polar Year, the schooner Tara drifted across the Arctic, Nansen-style, as part of the Developing Arctic Modelling and Observing Capabilities for Long-term Environmental Studies (DAMOCLES) project (Gascard et al. 2008, 2015); but its yearlong AABL program suffered from winter data gaps because of damage to instrumentation from ice deformation. During spring 2008, scientists flew in and collected boundary layer profiles with a tethered balloon and surface-flux observations (e.g., Vihma et al. 2008).
energy budget, compared to lower latitudes. For large parts of winter, the sun is absent; in summer it is above the horizon all day, but at large solar zenith angles and often in combination with a high surface albedo. Hence, there is typically a much-reduced diurnal cycle in the AABL, but a very strong annual cycle (Persson et al. 2002); see Fig. 9-27. This leaves a proportionally larger role for longwave radiation processes in determining the AABL stability and structure, and also provides a favorable setting for persistent conditions; low-level clouds, strongly modulating both short- and longwave radiation, are frequent in the Arctic.

The second factor is the thermodynamics of phase change (Persson 2012; Persson and Vihma 2017). In winter, the snow surface in the Arctic can become quite cold, especially in cloud-free conditions, while the ocean beneath the sea ice remains at the freezing point, \(-1.8^\circ C\). Heat conduction through the ice, from the ocean to the atmosphere, keeps the ice surface warmer than the adjacent land, while heat loss at the surface is translated to growing ice thickness since the temperature at the bottom of the ice remains at the freezing point. In spring and summer, solar radiation raises the surface temperature but only to the melting point, where it stays as long as significant amounts of melting snow and ice remain. Hence there is little or no diurnal cycle in summer and winter. Only when solar radiation is present and the surface temperature is well below the melting point can there be any diurnal cycle; even then it is not very pronounced. In the absence of nearby open water, the static stability of the AABL over ice is instead determined by clouds or advection of warm air from lower latitudes.

1) SURFACE-LAYER TURBULENCE

The first attempts to determine the surface momentum flux over sea ice were aimed at understanding how winds force the motion and deformation of the sea ice. Early investigators measured wind speed profiles on homogeneous sea ice, using similarity theory to determine roughness length \(z_0\) or the drag coefficient \(C_D\). Although low masts were used (<2 m; Untersteiner and Badgley 1965) to ensure a small footprint to more easily sample homogeneous conditions, they still found large variability, with \(z_0 \sim 10^{-5} - 10^{-3} m\). Already during AIDJEX, it was realized that large part of the atmospheric forcing came from form drag associated with pressure ridges in the ice (e.g., Arya 1973). This has since been verified from several field experiments [e.g., Arctic Ocean Experiment 2001 (AOE-2001); Tjernström 2005]. Momentum flux from the atmosphere contributes to such deformation of the ice, thereby modifying the surface drag, in turn changing the forcing. Roughness length is also influenced by the fraction of ice upstream as illustrated by Arctic Summer Cloud Ocean Study (ASCOS) observations from the edge of an open lead (Fig. 9-28). Note how \(z_0\) varies for different sectors, being smallest for homogeneous sea ice or open water fetch, but orders of magnitude larger for fetch over broken ice. Flux estimates using measurements from more complex locations, by open leads, and from moving platforms (ships and aircraft) have established that \(z_0\) for partial ice cover has a strongly nonlinear relationship with sea ice concentration, with the largest \(z_0\) for 40%–70% ice cover and the reason is—again—form drag, this time coming from many small ice edges between the ice floes (e.g., Schröder et al. 2003; Lüpkes et al. 2012).

Very special conditions develop when a lead in the ice opens up in winter. Open water at the freezing point (~\(-1.8^\circ C\)) is then exposed to air temperatures easily several tens of degrees below zero, leading to dramatically large surface heat fluxes (e.g., Andreas 1980;
Morison et al. (1993; Ruffieux et al. 1995), followed by a rapid formation of new ice. This process was a focus of some field campaigns (e.g., LEADEX); modeling (e.g., Mauritsen et al. 2005; Lüpkes et al. 2008) suggests an AABL impact far downwind. Aside from leads, early interest for turbulent heat fluxes was small, but this changed with a growing interest in climate change and its impact on sea ice. Thus SHEBA was designed to address the surface energy budget with turbulent surface heat fluxes as an important component. These fluxes were found to be typically small over sea ice: Fig. 9-29a shows sensible heat flux was most often close to zero, varying between $-30$ and $+20 \text{ W m}^{-2}$, while the latent heat flux was even smaller in magnitude and rarely downward. However, the frequency and magnitude of downward latent heat flux may be underestimated since ice is more likely to form on the sensors, which disrupts the observations under conditions with downward water vapor transport.

2) ARCTIC BOUNDARY LAYER CLOUDS AND SURFACE RADIATION

Surface and ABL energy budgets are driven by radiation; the main modulator for both shortwave and longwave radiation in the Arctic is cloudiness. Clouds are ubiquitous in the Arctic; the summer cloud fraction...
is typically >90%, dropping to 60%–80% in winter, and most of the clouds are AABL clouds (Shupe 2011).

The first studies of Arctic clouds came as spin-off from research flights during AIDJEX and focused on radiation (e.g., Herman 1977), followed by more focused aircraft campaigns, such as the 1980 Arctic Stratus Experiment (e.g., Curry and Herman 1985). Modeling came into play relatively early and one focus was the observed “layering” of Arctic clouds. Herman and Goody (1976) used model calculations of AABL clouds to infer the physics behind this layering, while manipulating turbulent mixing and the cloud’s radiative properties. As most of these earlier studies were either directly or indirectly based on observations from research aircraft, focus was on the summer, and clouds were generally assumed to contain only liquid water. Curry (1986), who provided an important review of the relevant processes and their interactions, was one of the first to point out the moisture inversion at cloud top, which leads to entrainment of moist air, and thus contributes to cloud maintenance and high AABL relative humidity, usually close to 100% in summer.

The study of AABL clouds took a giant leap forward with SHEBA along with a supplementary aircraft campaign targeting clouds and aerosols in spring 1998 (FIRE-Arctic; Curry et al. 2000). Perhaps the most important result was the observation that AABL clouds mainly consist of a mixture of ice particles and liquid droplets even in summer, and that a liquid-droplet layer can remain with cloud-top temperatures as low as $-34^\circ\text{C}$ (Intrieri et al. 2002b). This has a tremendous impact on the AABL energy budget and on its vertical structure. In contrast to ice clouds, which interact only weakly with infrared radiation, a liquid-water layer keeps the net surface longwave radiation near zero during overcast conditions; otherwise it is large and negative (upward). The transition between cloud-free or optically thin clouds and radiatively active clouds is often abrupt, tied to intrusions of warm and moist air from the south (e.g., Persson et al. 2002; Morrison et al. 2011). Stramler et al. (2011) showed how this leads to the now famous bimodal distribution of the surface net longwave radiation in Fig. 9-29b. With a sufficiently high surface albedo and large solar zenith angle, data taken during the summer (ASCOS) indicate that clouds can increase the surface net longwave radiation more than they reduce the net surface solar radiation (e.g., Sedlar et al. 2011). Intrieri et al. (2002a) showed that AABL clouds during SHEBA had a net warming effect at the surface most of the year, except for a short period in summer when the surface albedo was reduced; this effect is very sensitive to surface albedo (Sedlar et al. 2011; Intrieri et al. 2002b).

Clouds thus have a tremendous impact on the stability of the AABL, and hence on its surface turbulent fluxes and its vertical structure, as shown in Fig. 9-30. In essence, AABL clouds move the large upward surface net longwave radiation to the cloud top, where it generates cloud-top cooling, negative buoyancy, and hence turbulent mixing, as in the subtropical stratocumulus described in section 8 but without a substantial diurnal cycle (e.g., Tjernström 2007). This generates a cloud mixed layer, which often extends below the cloud.
Cloud-driven mixing appears to be a main factor for mixing the AABL, often more efficient than the surface mixing (Brooks et al. 2017). The cloudy turbulent layer is often decoupled from the surface (Shupe et al. 2013; Sotiropoulou et al. 2014; Brooks et al. 2017); limited observations, mostly from summer, indicate that this happens about two-thirds of the time.

AABL mixed-phase clouds are characterized by a thin supercooled liquid layer that constantly precipitates ice crystals. This has generated substantial interest, both in a cloud microphysical context and for its effects on AABL vertical structure and stability. A series of field campaigns was devoted to AABL mixed-phase clouds: FIRE-Arctic (Curry et al. 2000), Mixed-Phase Arctic Cloud Experiment (M-PACE; Verlinde et al. 2007), and Indirect and Semi-Direct Aerosol Campaign (ISDAC; McFarquhar et al. 2011). Although many of these campaigns were primarily focused on cloud microphysics, the results confirm that cloud-driven turbulent mixing in the AABL plays a crucial role in maintaining both cloudiness and AABL turbulence, as summarized in a review paper by Morrison et al. (2011). Sublimation/evaporation of ice particles/drizzle from the cloud layer and evaporation from the surface may saturate the lower layer; in that case fog often develops.

3) VERTICAL STRUCTURE OF THE AABL

The scarce early observations suggested that the Arctic AABL is usually stably stratified; but is this correct? Already Sverdrup (1925), while analyzing the temperature profiles from the Maud expedition, concluded that the temperature structure over the Arctic Ocean was different from that over the continent in that “every wind generally also brings an increase in temperature with height.” Since then many papers have discussed various aspects of the “stable Arctic boundary layer” without distinguishing whether the inversion is elevated or at the surface. Some, however, seem to have understood this distinction. In the 1950s, Webster (1954, p. 77), in a review of measurements from the Russian ice-drift stations, questions Sverdrup and writes, “The lower boundary of the layer of inversion is marked by turbulence which causes a vertical distribution of temperatures characteristic of unstable air masses. The so-called ‘cold layer’ would be better named the ‘unstable layer.’” Although the word “unstable” is perhaps unfortunate in a boundary layer context, since the AABL is seldom truly convective, it is clear that he was referring to the shallow well-mixed inversion-capped AABL. Interestingly, Webster also discussed the advection of warm air into the Arctic: “These masses often disperse above the 100–200 m cold earth layer throughout the troposphere to a height of 7–8 km.” Although not strictly a boundary layer-meteorology feature, and hence not further discussed here, it is notable that this advection contributes to maintaining and strengthening the deep AABL-capping inversion and may also explain the moisture inversion. It is also among the most discussed issues in Arctic meteorology today.

Several papers in the early 1990s used soundings from combinations of coastal stations, aircraft dropsondes, and the Russian NP stations (Kahl et al. 1999) to explore the stratification in the Arctic lower troposphere, and concluded that inversions were frequent (Kahl 1990; Serreze et al. 1992) and their strength was increasing (Kahl et al. 1996). Based on the year of high-resolution SHEBA soundings, Tjernström and Graversen (2009) showed that a lower-troposphere inversion was almost always present. Considering only boundary layer inversions, they found that the mean inversion base height was typically 100–400 m from April through September, while surface inversions were most common from December through February. These inversions were commonly 500–700 m thick from November through March, and thinner (200–400 m) though much of spring and summer (Fig. 9-31). While surface inversions were present less than 10% of the time during spring and summer, they occurred slightly over half the time during autumn and winter, making the dominant AABL stratification during SHEBA nearly neutral.

Based on the foregoing, the main reason for the predominantly well-mixed AABL structure is the presence of AABL clouds. Whenever these are present, the combination of cloud-driven mixing and mechanical mixing at the surface keep the AABL nearly neutrally stratified, in summer as well as in winter. When the clouds disappear, the surface energy budget is dominated by strong negative net longwave radiation flux, the temperature drops, and a surface inversion forms. This effect is most pronounced in winter, when the presence of clouds seems to be governed by large-scale advection, explaining the persistence of both regimes and the rapid transitions between them (Morrison et al. 2011). During the SHEBA winter, the free troposphere temperature was systematically higher in cloudy than in clear conditions (O. Persson 2018, personal communication). Unlike in the nocturnal continental SBL in lower latitudes, the stably stratified AABL is rarely capped by a near neutrally stratified residual layer, so that vertically propagating gravity waves in the free troposphere can reach the surface (Perssson and Vihma 2017), where they can break and contribute to AABL turbulence. In summer, however, AABL clouds are common and persistent and therefore surface inversions are rare. Hence the conceptual model in Fig. 9-30 can serve as a
zeroth-order description of the AABL, almost all the time in summer and almost half the time in winter.

c. Modeling and the future

Given the lack of initialization and verification data, our still emerging knowledge of the dominating physics, and the strong feedbacks between clouds and fluxes, it is not surprising that numerical models have difficulty replicating the complex behavior of the AABL. Sorteberg et al. (2007) tried to make up for the data deficiency by comparing surface fluxes in climate models to those from reanalysis, essentially performing a model-to-model intercomparison. Comparing with actual data, Tjernström et al. (2005) found that models did poorly in replicating sensible and latent heat fluxes from SHEBA, with too many extreme values of both signs. Momentum flux was better, but with too many cases with large values and too few with small. Similarly, Tjernström et al. (2008) and Wyser et al. (2008) showed that regional climate models systematically produced winter AABL clouds that were too optically thin (mostly being ice clouds), and summer clouds that were too optically thick. Consequently, surface net longwave radiation flux was much too small in winter, while surface solar forcing was too low in summer (e.g., Prenni et al. 2007). Besides affecting the AABL itself, all these biases have severe consequences in coupled modeling because of the sensitivity of the sea ice to atmospheric forcing.

10. Representing the ABL in models

This section describes how the knowledge gained over the last 100 years and earlier has been synthesized into atmospheric dispersion models, land surface models, and ABL parameterization schemes. Dispersion models provide us with tools for predicting the spread and concentration of contaminants. Land surface models account for the strong influence of the surface and its heterogeneity on the ABL, while ABL parameterization schemes account for boundary layer effects in weather and climate models.

a. Dispersion models

Progress in dispersion understanding is paced by our knowledge of the ABL and its turbulence. Here, we review the key dispersion-modeling developments in mostly chronological order for 1) the early period (prior to 1960–70), 2) the “explosive period” (from 1970 to ~1990) with main achievements for convective and stable ABLS, and 3) progress on other topics (from ~1990 to present)—(i) LES of dispersion, (ii) Lagrangian particle models, (iii) concentration fluctuations, and (iv) urban dispersion.

1) EARLY PERIOD: EDDY DIFFUSION, LAGRAGIAN SIMILARITY, AND STATISTICAL THEORIES

Early dispersion research was restricted to heights less than ~100 m because of limited turbulence and wind information. Dispersion was modeled using eddy-diffusion methods, Lagrangian similarity theories (LSTs), statistical theories (e.g., Pasquill and Smith 1983), and analytical methods. In principle, eddy diffusion or “K” theory is limited to problems where the local plume depth (e.g., vertical dispersion $\sigma_z$) is of the same order or larger than the characteristic eddy size (e.g., Weil 2013). Short-range dispersion from a surface...
release satisfied this limitation, and the two-dimensional diffusion equation with an appropriate $K$ was used to model the crosswind-integrated concentration (CWIC $C'$) from a surface source. This was done with Roberts’s (unpublished) analytical solution (Calder 1949) and power-law profiles of wind and diffusivity, and achieved partial agreement with observations. Vertical dispersion also was modeled with LST (Monin 1959; Gifford 1962), which assumed that the Lagrangian velocity of “particles” depends on the same variables, which assumed that the Lagrangian velocity of “particles” depends on the same variables, $u_*, \theta_*$, buoyancy, etc., as given by the Eulerian statistics and MOST. However, LST gave no real advantage over $K$ theory and has seen little use since the 1980s.

For statistical theory, two key developments took place both for homogeneous turbulence and based on Lagrangian formulations: 1) Taylor’s (1921) results on mean dispersion in a fixed (Eulerian) system, and 2) Batchelor’s (1950) theory on relative dispersion, that is, dispersion about the local plume centerline [see section 10a(3)(iii)]. Taylor’s analysis predicted the root-mean-square (rms) vertical dispersion to vary as $\sigma_z = \sigma_w t$ for short times $t < T_L$ and as $\sigma_z = (2\sigma^2 w T_L t)^{1/2}$ for long times $t > T_L$, where $\sigma_w$ is the standard deviation of $w$, and $T_L$ is the Lagrangian integral time scale or “memory time” of the dispersion process. The short-time result is sometimes called a “ballistic” limit because the particle paths are essentially straight lines, whereas the long-time result is consistent with $K$ theory, that is, a diffusive ($t^{1/3}$) limit or “random walk,” with a vertical diffusivity $K_z = \sigma^2_w T_L$. Similar results apply to other directions. This theory has withstood the test of time, being relevant today in providing 1) theoretical limits for new models, 2) dispersion predictions for convective and stable ABLs, and 3) expressions for applied dispersion models based on parameterized turbulence profiles (Hanna and Paine 1989; Carruthers et al. 1994; Cimorelli et al. 2005; see also Hupt et al. 2019).

2) EXPLOSIVE PERIOD: CONVECTIVE AND STABLE BOUNDARY LAYERS

The “explosive period” is so labeled because of the rapid and significant progress in the understanding of ABL turbulence, structure, and dispersion that was achieved through laboratory experiments, numerical simulations, and field observations. For dispersion from surface sources, Nieuwstadt and van Ulden (1978) extended the “early period” work by solving the diffusion equation numerically using MOST profiles for mean wind and turbulence, and van Ulden (1978) obtained an analytical solution based on an extension of Roberts’s work to include the MOST profiles. These models were applied to both convective and stable surface layers using the MOST similarity function for heat $\phi_h(z/L)$ (section 3b). CWIC results from the models agreed well with the Prairie Grass data (Barad 1958), which are considered the “gold standard” for surface releases. Related modeling was conducted by Gryning et al. (1983) for surface deposition and Horst (1979).

Venkatram (1992) provided further understanding by deriving simple analytical expressions for $\sigma_z$ and $\sigma_u$ for the surface CWIC in very unstable and stable conditions using the long-time limit of statistical theory $\sigma^2_w/dt = K_z(t)$ and the MOST results for heat, $K_z = K_{hi}(z/L)$. By evaluating $K_{hi}$ at $z = \sigma_z$, he obtained $\sigma_z \approx x^2/L$ and the surface $C' \approx Q_S(L)/(u_\infty x^2)$, in very unstable conditions, with $\sigma_z \approx L^{2/3}x^{1/3}$ and $C' \approx Q_S(u_\infty L^{1/3}x^{2/3})$ in very stable conditions, where $Q_S$ is source strength and $x$ is the downwind distance. The downwind dependence on $x$ in Vankatram’s expressions matched van Ulden’s (1978) curves and explicitly showed the variation of $\sigma_z$ and $C'$ with $x$, $L_z$, and $u_\infty$.

(i) CBL

Convective scaling (Willis and Deardorff 1976) was a major advance in understanding of dispersion throughout the CBL and enabled the comparison of laboratory, numerical, and field-data results in simple dimensionless coordinates. The key length, velocity, and time scales for turbulence and dispersion were $z_i$, $w_\infty$, and the eddy turnover time $z_i/w_\infty$, where $T_L = z_i/w_\infty$ for the large eddies. In strong convection ($-z/L > 1$), the mean wind $\overline{U}$ is nearly uniform in the upper CBL and thus the plume travel time $t = x/\overline{U}$. Willis and Deardorff proposed that the dimensionless CWIC, $C'/\overline{U}z_i/Q_S = f_1(z_i/z_s, X, z_i/z_s)$, where $X = w_\infty x/(\overline{U}z_i)$ is the dimensionless distance or travel time, $z_s$ is the source height, and $f_1$ is some functional form. Here, $C'/\overline{U}z_i/Q_S$ is the ratio of the local CWIC $C'$ to the well-mixed value $Q_S/\overline{U}z_i$ attained far downstream ($X \sim 3$).

Figure 9-32 shows contours of the dimensionless CWIC from the Willis and Deardorff (1976, 1978) laboratory experiments as a function of dimensionless height and distance in strong convection ($-z/L > 1$). For a near-surface source, the average plume centerline (as defined by maximum CWIC) ascends after a short distance, whereas an elevated source centerline descends until it reaches the surface. These behaviors were explained by the vertical-velocity skewness and the release of material into updrafts and downdrafts (Lamb 1982; Weil 1988), which have different areal fractions. More area (60%) is covered by downdrafts, and thus for the elevated source, there is a higher probability of material being emitted into downdrafts and then being carried down toward the surface. The results have been supported by field observations (Nieuwstadt 1980; Briggs 1993) and numerical simulations (e.g., Lamb
1978, 1982; Baerentsen and Berkowicz 1984; Weil et al. 2012). They differed substantially from the standard Gaussian plume model of the time, which predicted a horizontal centerline.

The elevated source experiment referred to above was motivated by Lamb’s (1978) numerical simulation using a Lagrangian particle dispersion model (LPDM) driven by LESs. Lamb (1978) first predicted the descending plume centerline for the elevated release, which was later confirmed by experiment (Fig. 9-32b) and field observations (Briggs 1993).

Comparisons of statistical theory with observations and convective scaling have been made for elevated sources in the CBL (Nieuwstadt 1980; Briggs 1988, 1993; Weil 1988); the theory is applicable because the mean wind and turbulence are quasi-homogeneous in the upper CBL. Figure 9-33a shows that crosswind dispersion (σ_<sub>y</sub>) observed from the Convective Diffusion Observed with Remote Sensors (CONDORS) experiment (Briggs 1993) in dimensionless form agree with the statistical theory prediction, σ_<sub>y</sub>/z<sub>i</sub> = 0.6X, which holds for X < ~ 0.7; beyond this X, the observations suggest that long-time (t > T<sub>L</sub>) memory effects may be important with σ_<sub>y</sub> ∝ t<sup>1/2</sup>.

For vertical dispersion (Fig. 9-33b), the data support the short-time prediction for X < 0.5, but beyond that distance, they converge toward σ_<sub>z</sub>/z<sub>i</sub> = 0.27. The latter is due to plume trapping by the elevated inversion and a tendency toward a vertically well-mixed distribution by X ~ 2 or 3. Results from the Willis and Deardorff (1976, 1978, 1981) experiments and Lamb’s numerical simulations (Lamb 1978, 1982) are similar to the observations.

These results motivated a new type of dispersion model—an approach based on the probability density function (PDF) of w in the CBL (Misra 1982; Venkatram 1983; Weil 1988). The PDF is non-Gaussian, consistent with the Fig. 9-32 results, and is typically modeled with a bi-Gaussian distribution. It has been included in applied dispersion models (e.g., Cimorelli et al. 2005; Carruthers et al. 1994; Hanna and Paine 1989).

The above results pertained to strong convection, but data were sorely needed for the more common weaker convective regime (z<sub>i</sub>/L ≪ ~ 20). Dosio et al. (2003)
used LES to explore dispersion in this regime and found reduced vertical dispersion with lower mean plume heights versus X, surface CWIC versus X profiles shifted downstream, and increased lateral dispersion ($\sigma_z$); this was attributed to the greater TKE dissipation rate near the surface and shorter time scales $T_L$ for weaker convection. Another extension of the CBL work was to buoyant plume dispersion, pursued in part using LES (van Haren and Nieuwstadt 1989), but mostly explored in convection tank experiments (Willis and Deardorff 1983, 1987; Weil et al. 2002). The experiments provided new insights into highly buoyant plumes that loft at the CBL top and resist downward mixing, and exhibit an enhanced lateral spread because of their buoyancy. The enhanced lateral spreading was successfully predicted using a gravity-current model (e.g., Simpson 1982).

(ii) SBL

For elevated sources in the SBL, the short-time vertical dispersion is given by statistical theory, but the long-time dispersion can range between a classical diffusion limit ($\sigma_z \propto t^{1/2}$; Venkatram et al. 1984) and a constant vertical thickness or “pancake limit” (Pearson et al. 1983); the latter was observed by Hilst and Simpson (1958). The “statistical theory” case was modeled using an interpolation expression of the form, 

$$\sigma_z = \sigma_u t/(1 + 0.5t/T_L)^{1/2},$$

which satisfied both the short and long time limits. In addition, Venkatram et al. (1984) parameterized $T_L$ using “neutral” and “stable” forms applicable near the surface and in the upper SBL, respectively, as given by Brost and Wyngaard (1978). Their model was supported by dispersion measurements downwind of a ~50-m tall source.

Pearson et al. (1983) predicted the long-time dispersion to range from a constant $\sigma_z$ to the classical behavior ($\sigma_z \propto t^{1/2}$) depending on the value of a “molecular mixing” parameter $\gamma$. For zero or small $\gamma$, a constant pancake-like plume is attained with $\sigma_z \propto \sigma_u/N$, where the Brunt–Väisälä frequency $N$ depends on the local stratification. That is, if $\gamma = 0$, a fluid particle moving about in a stratified turbulent flow retains its initial density and moves through a vertical distance $\sim \sigma_u/N$ before the buoyancy force returns it to its original height. However, if $\gamma > 0$, the particle can be warmed or cooled by molecular diffusion if moving up or down, weakening the buoyancy force; the particle is then displaced a distance greater than $\sigma_u/N$. With a moderate to large $\gamma$, the classical behavior is recovered. The Pearson et al. (1983) model was supported by observations, including field data in the Venkatram et al. (1984) study (see Weil 2013).

The classical dispersion limit would be expected in a very stable boundary layer with weak, intermittent or “wavylike” turbulence in the upper SBL. Further support for the classical $\sigma_z$ behavior was given by dispersion simulations using the coupled LPDM-LES for a WSBL (Kemp and Thomson 1996). Recent LESs of a moderately stable boundary layer have been conducted for fine grid spacings (0.39–12.5 m) (Beare et al. 2006; Huang and Bou-Zeid 2013; Sullivan et al. 2016), with those by Sullivan et al. conducted for the finest grid spacings. Their results for $z_i/L = 6$ show a region of relatively weak turbulence ($\sigma_u \approx 0.01$ m s$^{-1}$) in the uppermost 30% of the SBL, which offers the potential for finding the pancake-like dispersion behavior using an LPDM driven by LES.

Lateral dispersion can be determined from statistical theory, but parameterization of $\sigma_u$ in terms of the micrometeorological variables $u_*, L$, etc. is problematic because of contributions from large-scale meandering, especially in weak winds. Mahrt et al. (2001) found that the source of meandering was ill defined. Hence, Vickers et al. (2008) used data from a 7-tower network over a small domain to measure the spatially and temporally varying wind field along with an added stochastic component to predict “particle” dispersion and interesting horizontal concentration contour patterns. The latter consist of “spokelike” or “jagged” crosswind concentration profiles because of localized peaks of wind-direction probability rather than a smoothly varying wind-direction probability. Similar work was done by Anfossi et al. (2006), who compared their concentration results with observations. One potential source of the meander could be the ubiquitous random turbulent structures and fronts in the SBL surface layer (Adrian 2007; Sullivan et al. 2016).

3) PROGRESS ON OTHER TOPICS

Over the past three decades, large-eddy simulation has gained much use in dispersion modeling mainly because of its success in ABL research. Here, we discuss this LES application as well as Lagrangian particle modeling, concentration fluctuations, and urban dispersion.

(i) LES of dispersion

The application of LES to dispersion modeling has followed two main approaches: 1) Eulerian methods in which the scalar diffusion equation is added to the LES model, and 2) Lagrangian techniques wherein “passive” particles from a source are tracked using the LES velocity fields [see sections 10a(2) and 10a(3)(ii)]. Here, the focus is on the Eulerian approach in which early contributions were made by Sykes and Henn (1992) and Henn and Sykes (1992) for neutral and convective ABLs, respectively. For mean concentrations, their
results were in fair-to-good agreement with experiments in a neutral wind-tunnel flow (Fackrell and Robins 1982) and a convection tank (Willis and Deardorff 1978, 1981). Their computed concentration fluctuations for the CBL, including variances and PDFs, were some of the first from an LES and demonstrated the LES potential. However, their modeled fluctuations were typically only 60% of the tank data values possibly because of the small number of grid cells (80 × 80 × 72), the large source size modeled, and the neglect of SGS fluctuations; these problems have been overcome with a Lagrangian two-particle model driven by LES [see Weil et al. 2017; section 10a(3)(ii)]. The Eulerian approach also was adopted in studies of CBL stability effects on dispersion [Dosio et al. 2003; see section 10a(2)] as well as dispersion statistics and concentration fluctuations (Dosio and de Arellano 2006), which matched the tank fluctuation data much better than Henn and Sykes (1992). This was likely due to the much greater number of grid cells (512 × 512 × 128) used by Dosio and de Arellano (2006) than by Henn and Sykes (1992).

In addition, the LES Eulerian approach has been applied to 1) pollen dispersion in a neutral ABL over a crop field (Chamecki et al. 2009), 2) fumigation of pollutants trapped in an elevated inversion into the CBL mixed layer (Cai and Luhar 2002), and 3) dispersion in a regular array of urban block buildings for both a neutral LES flow (Boppana et al. 2010; Tseng et al. 2006) and a stably stratified flow for a similar building configuration (Tomas et al. 2016).

(ii) LPDMs

Over the past 30 years, Lagrangian particle dispersion models have seen greater use because of 1) their capability of handling turbulence complexities such as spatial inhomogeneity, non-Gaussianity, and a large \( T_z \), and 2) the substantial increase in computing power. In an LPDM, one computes “passive” particle trajectories given the random velocity field with the mean concentration found from the particle position PDF, which is obtained by tracking tens to hundreds of thousands of particles. The velocities are determined either from 1) a purely stochastic approach—a Lagrangian stochastic model (LSM), or 2) LES using the resolved LES velocities plus a random subfilter-scale (SFS) velocity obtained from an LSM. In a landmark paper, Thomson (1987) gave a rigorous basis for an LSM construct, using two criteria: 1) the “well mixed” condition that assumes that particles initially uniformly distributed in phase space must remain so, and 2) the Fokker–Planck equation that governs the evolution of the joint PDF of particle position and velocity. Thomson’s approach corrected some problems of earlier models and has been adopted in most currently used LSMs. Some examples include dispersion in the CBL (Luhar and Britter 1989; Weil 1990), which requires a skewed \( w \) PDF, fumigation into a coastal internal boundary layer (Luhar and Sawford 1995), and dispersion from surface sources over a range of stability (Du and Venkatram 1997). Much of this work was reviewed by Wilson and Sawford (1997), who reported problems for strongly stratified turbulence, which remains a key issue for ABL research. The review led Das and Durbin (2005) to develop a new approach that improved the LSM agreement with experiments on stable flows.

In pioneering work, Lamb (1978, 1982) developed the LPDM-LES approach and simulated dispersion in the CBL, finding good agreement of dispersion statistics and concentration fields with the Willis and Deardorff (1976, 1978) tank experiments. Weil et al. (2004) adapted a more detailed LSM (Thomson 1987) for the SFS velocities and a finer LES grid and also obtained agreement with the Willis and Deardorff data; they also demonstrated the need for a realistic SFS model to predict surface concentrations for near-surface sources. Other ABL applications include 1) fumigation in a CBL (Kim et al. 2005), 2) dispersion in the SBL (Kemp and Thomson 1996), 3) dispersion in the evening transition layer (Taylor et al. 2014), and 4) statistical variability of dispersion in the CBL based on LPDM-LES modeling and observations (Weil et al. 2012).

The LSM and LPDM-LES methods also have been used for determining flux footprints—the contribution of upwind surface emission elements to the measured vertical scalar flux at some height. They have both a “forward” formulation from source to receptor and a “backward” version. LSM applications for the forward case include 1) the atmospheric surface layer (Horst and Weil 1992; Leclerc and Thurtell 1990; Schmid 1994), 2) the CBL (Weil and Horst 1992; Kljun et al. 2004), and 3) flows within and above canopies (Baldocchi 1997). The backward LSM has the advantage of being applicable to nonhomogeneous surfaces and has been used by Flesch et al. (1995) and Kljun et al. (2002, 2004).

(iii) Concentration fluctuations

Concentration fluctuations result from the stochastic nature of turbulence and dispersion, and are important for estimating the concentration variance, PDF, and peak values (Csanady 1973; Chatwin 1982). The fluctuations at a point are characterized by \( \sigma_c \) or the fluctuation intensity, \( \sigma_c/C \), where \( C \) is the ensemble mean concentration at that point; for an elevated source the intensity can range from 1 to 10 at short downwind distances, with the high values caused by large-scale plume meandering. Several approaches have been pursued for predicting the fluctuations and \( \sigma_c \): 1) models
based on Gifford’s (1959) “meandering plume” approach and extensions using LSMs (Franzese et al. 2003; Luhar et al. 2000), 2) Lagrangian two-particle stochastic models that predict the relative dispersion and $\sigma_\epsilon$ (Thomson 1990), 3) micromixing models that describe the time variation of a “particle’s” concentration with coupling to LPDMs for concentration statistics (Sawford 2004; Cassiani et al. 2005; Luhar and Sawford 2005), 4) second-order closure models (Sykes et al. 1984), and 5) LES coupled with the diffusion equation (Dosio and de Arellano 2006; Henn and Sykes 1992; Sykes and Henn 1992) and more recently LES coupled with a Lagrangian two-particle dispersion model (L2PDM; Weil et al. 2017).

The above approaches either explicitly model the relative dispersion $\sigma_r$ about the plume centroid at a given distance from its source with Batchelor’s (1950) theory or use the theory for estimating key variables such as the micromixing time scale (e.g., Cassiani et al. 2005). For short times, $\sigma_r \propto t^{1/2}$ and at long times, $\sigma_r \propto t^{1/2}$. The models have been tested with laboratory data for the CBL (e.g., Deardorff and Willis 1984; Weil et al. 2002) and a neutral boundary layer (Fackrell and Robins 1982).

(iv) Urban dispersion

Understanding of dispersion in urban areas has progressed substantially over the past 25 years through the combined work of modeling and simulations, laboratory experiments, and field observations. Urban dispersion is often classified in 1) four horizontal-scale regimes—regional ($x < 200$ km), city ($< 20$ km), neighborhood ($x < 2$ km), and street ($x < 200$ m); and 2) four vertical sublayers analogous to those considered in section 3—urban canopy layer (UCL; $z/h_B < 1$), RSL ($z/h_B < 2$ or 3), inertial sublayer ($z/h_B > 2$ or 3), and outer layer; here $h_B$ is the mean building height. Recent reviews by, for example, Britter and Hanna (2003), Belcher (2005), Fernando (2010), and Belcher et al. (2013) discuss this progress in detail. They show that the coordinated efforts of theoretical and numerical modeling (e.g., Coceal et al. 2014; Tomas et al. 2016), wind-tunnel studies (e.g., Kastner-Klein et al. 2004; Castro et al. 2006), and field observations (e.g., Davidson et al. 1995; Hanna et al. 2007) have improved knowledge. The urban flow field is marked by increased turbulence, mainly in the UCL and RSL, because of flow about the buildings and surface warming, while the mean wind is reduced because of building drag. Vertically, the effects are greatest in the UCL and RSL, and horizontally on the neighborhood and street scales. Studies show that 1) the locations of shear layers and “dividing streamlines” in the UCL control the mixing and transport (Belcher 2005), and 2) lateral dispersion varies initially as $\sigma_x \propto x$ and then as $\sigma_x \propto x^{1/2}$ because of “topological dispersion” or wind-direction changes from flow impingement on buildings in an array (Davidson et al. 1995; Jerram et al. 1995). In addition, an isolated tall building in an urban area greatly enhances the upward transport of polluted UCL air because of the low pressure and corner vortices on the building lee side (Heist et al. 2009).

An unresolved issue is the net effect of urban roughness on maximum concentrations from surface sources due to the competing effects of increased turbulence and reduced winds, which lead to lower and higher concentrations, respectively. Some small-scale field and laboratory experiments show little net effect (Britter and Hanna 2003), whereas field observations and modeling demonstrate a net concentration reduction (Britter and Hanna 2003; Delle Monache et al. 2009).

b. Land surface modeling in weather and climate models

To a large degree, the ABL is driven by surface fluxes, which in turn result from the exchange of energy and water fluxes across the land (or water)–atmosphere interface, as described in section 3a. The exchanges over the oceans are described in section 3d. Here are described the much more complex exchanges over land, as represented in land surface models (LSMs), which follow the flow of moisture and heat from the subsurface, through the surface, its vegetation, (and sometimes built structure), and snow/ice into the atmospheric surface layer, where the fluxes are linked to ABL parameterization schemes through MOST functions. Land surface processes play an important role in both GCMs (e.g., Mintz 1982; Rowntree and Bolton 1983), and regional and mesoscale atmospheric models (Ookouchi et al. 1984; Mahfouf et al. 1987; Chen and Avisar 1994a, b; Pielke et al. 1997; Rasmussen et al. 2011).

In NWP and climate models prior to the 1960s, the description of land surface processes was either absent or oversimplified with a prescribed diurnal cycle of surface heat fluxes and surface temperature. The first interactive model concept was perhaps the work of Budyko (1956), which used a simple “bucket” model to calculate surface evaporation as a function of soil moisture. That concept was perhaps first adapted for use in a climate model by Manabe (1969) and then for an interactive soil model in a mesoscale weather model by Physick (1976). McCumber and Pielke (1981) coupled a mesoscale model with a soil model in which the surface heat fluxes depended on the soil moisture and temperature.

Undoubtedly, introducing a vegetation foliage layer (the so-called big leaf model) by Deardorff (1978) was a major milestone in land surface modeling. The concept
of explicitly treating the plant canopy with simplifications was later adopted in land surface models (e.g., Pan and Mahrt 1987; Noilhan and Planton 1989) or enhanced with more complex biophysical processes (Dickinson 1984; Sellers et al. 1986) and hydrological processes (e.g., Entekhabi and Eagleson 1989; Wood et al. 1992).

In the 1990s, the availability of long-term field data such as HAPEX-MOBILHY (André et al. 1986), Cabauw (Beljaars and Bosveld 1997) and FIFE (Sellers et al. 1986) promoted further development of LSMS, and systematic evaluations through initiatives such as the Project for the Intercomparison of Land-Surface Parameterization Schemes (PILPS; Henderson-Sellers et al. 1993; Wood et al. 1998), and their implementation in NWP models (e.g., Noilhan and Planton 1989; Chen et al. 1997; Smirnova et al. 1997; Viterbo and Beljaars 1995; Chen and Dudhia 2001). Figure 9-34 shows processes included in the Noah model, typical for an LSM used in NWP models.

In the context of the longer-time-scale Earth system models, such as the NCAR Earth System model and its LSM, the Community Land Model (CLM; Oleson et al. 2013; Lawrence et al. 2012), development today deals in greater depth with much broader topics, such as the spatial distribution of soil characteristics, vegetation, the use of remote sensing data, and the initialization of soil moisture and snow depth. This includes human impacts, such as urbanization (Best 2005; Oleson et al. 2008; Chen et al. 2011) and agricultural management (Levis et al. 2012; Liu et al. 2016; Leng et al. 2017). Developers continue to add more complexity and physical realism to LSMS, while starting to explore incorporating horizontal complexities such as subgrid-scale land-cover heterogeneity for large-scale models (e.g., Oleson et al. 2010; Li et al. 2013) and applying subgrid-scale methods developed for LES to heterogeneous surfaces at larger scales (Passalacqua et al. 2006). In particular, more LSMS include lateral surface and subsurface hydrologic flows at the grid or subgrid scales (e.g., Seuffert et al. 2002; Maxwell and Miller 2005; Miguez-Macho et al. 2007; Gochis et al. 2013; Barlage et al. 2015), additions that have had major impact on simulated surface heat and moisture fluxes, regional weather, and climate.
This increased complexity invites more uncertainty because of our incomplete knowledge of the physics, the inherent assumptions in model structures and parameterization schemes (Rosero et al. 2009, etc.), and the large set of imperfectly specified physical parameters used (Chen and Dudhia 2001). Mitigating these uncertainties requires more systematic evaluation and benchmarking of LSMs (e.g., Koster et al. 2004; Luo et al. 2012; Best et al. 2015; Newman et al. 2017). Thus, new frameworks have been developed for testing working hypotheses and parameterizations (e.g., Clark et al. 2011). The community Noah LSM with multi-parameterization options (Noah-MP; Niu et al. 2011) and the Structure for Unifying Multiple Modeling Alternatives (SUMMA; Clark et al. 2015) LSM are two prominent examples. Recent studies (e.g., Zhang et al. 2016) illustrate how incorporation of different parameterization schemes for key physical processes provides the opportunity to interpret modeling results for a given subprocess through the use of physics-ensemble simulations in the same model framework.

Last, it would not make much sense to implement a sophisticated LSM in weather models without a proper soil-state initialization procedure. However, there are no routine soil moisture and soil temperature observations over the large spatial domains typically spanned by weather models. Moreover, LSMs may have different soil moisture dynamic ranges because of different approaches to treating evaporation and runoff (Koster and Milly 1997), and the soil types used as input are only approximate. It is therefore important to recognize that the soil state calculated for a given atmospheric forcing depends on the model and input. The soil moisture and temperature fields from the traditional four-dimensional data assimilation (4DDA) systems of coupled land/atmosphere models often suffer substantial errors and drift owing to precipitation, temperature, and radiation biases in the coupled system. Such errors and drift often necessitate the application of soil moisture nudging techniques in coupled 4DDA systems (Mahfouf 1991; Douville et al. 2000). But the nudging approach may fail in situations where the soil moisture is insensitive to atmospheric conditions such as under cloudy and windy conditions. The combination of remotely sensed surface temperature and soil moisture, and the development of ensemble Kalman filter and variational methods enable large-scale land data assimilation (Maggioni and Houser 2017).

An alternative strategy is to utilize available observations (precipitation, surface solar insolation, and meteorological analysis) to drive an LSM in offline mode and use the results to initiate the model. This need was met by the development of the North American Land Data Assimilation System (NLDAS; Mitchell et al. 2004), the Global Land Data Assimilation System (GLDAS; Rodell et al. 2004), and the high-resolution land data assimilation system (HRLDAS; Chen et al. 2007).

c. Evolution of ABL parameterization in weather and climate models

Since the 1970s, operational and research NWP and climate models have become increasingly sophisticated, and this has included the treatment of the ABL, especially under convective conditions. Early models, especially global models, had low vertical resolutions with the lowest layers several hundred meters thick, thus representing the boundary layer as a whole rather than a separate thin surface layer for which similarity theory could be used. These thicker layers required so-called bulk ABL parameterizations and bulk exchange coefficients to derive surface fluxes from lowest-level atmospheric fields. The methods were typically simple with empirical heat and momentum exchange coefficients, as used in early mesoscale models (e.g., Anthes and Warner 1978). Phillips (1986) had a bulk mixed-layer scheme to redistribute surface fluxes in a boundary layer of several levels in the National Meteorological Center’s (NMC’s) global and regional Nested Grid Model (NGM) and Regional Analysis and Forecast System (RAFS) models.

Already in the 1970s, boundary layers in models were represented with higher resolution and vertical diffusion with mixing length theory, using enhanced mixing lengths in the boundary layer to account for eddy scales larger than the vertical grid size. O’Brien’s (1970) \(K\) profile, a cubic function of \(z\), was adopted by Tapp and White (1976) in the Met Office mesoscale model, while Busch et al. (1976) had a prognostic length scale that had a similar shape and was used in early versions of the Pennsylvania State University mesoscale model (Anthes and Warner 1978). This latter model also adopted the Deardorff (1972b) countergradient term to allow vertical heat fluxes even in well-mixed conditions. This term was to mimic the effect of deep surface-based eddies that transported heat nonlocally, that is, in a way that was independent of the local vertical gradient, and it allowed neutral stratification to develop, while one without such a term would remain unstable and could never reach neutrality when driven by surface fluxes alone. Zhang and Anthes (1982) presented a successor to the Pennsylvania State University high-resolution ABL scheme, following the ideas of Blackadar (1979), which represented direct transport by thermals between the surface layer and all other CBL layers, an early example of a nonlocal mass-flux approach. It also included an entrained buoyancy flux equal to \(-0.2\) times the surface buoyancy flux.
In the 1970s and 1980s, turbulence modeling also increasingly turned to so-called higher-order closure, inspired by the seminal paper on the hierarchy of turbulence closures by Mellor and Yamada (1974). Computational power was increasing rapidly in the 1970s, so there was a general belief that the number of prognostic equations for turbulence statistics could be increased, thereby pushing uncertain closure assumptions to even higher moments, where they either would either be simplified or at least not be so critical. For example, André et al. (1978) presented a model with prognostic equations for both second- and third-order correlations, which was tested against the Wangara data. Eventually, however, it was realized that little was gained, the closure assumptions for higher moments were more complicated than expected, negative variances resulted (Wyngaard 1988), and results were poor for the CBL (Moeng and Wyngaard 1989). Moreover, some of the hybrid schemes, like the Mellor–Yamada Level 3 scheme, were not fully realizable (e.g., André 1990).

Nevertheless, the so-called 1.5-order schemes, which predicted TKE as an extra second-order variable, started to become popular in the early 1990s. Such a scheme was adopted for the regional Eta model at NMC (Janjic´ 1990) based on the ideas of Mellor and Yamada (1982). Together with a diagnosed vertical length scale, TKE could be used to compute the vertical K coefficients. A prognostic TKE approach could be applied to the whole column, handling elevated turbulence as well as boundary layer turbulence, while maintaining a memory of the turbulence even after surface forcing ceases. French global and regional NWP models [ARPEGE, ALADIN, Applications of Research to Operations at Mesoscale (AROME)], and the European HIRLAM model consortium also adopted such approaches (e.g., Bougeault and Lacarrere 1989; Cuxart et al. 2000); the various TKE approaches mostly differ by how they compute the length scales and stability effects. The Bougeault–Lacarrere approach also adopts Deardorff’s (1972b) countergradient term. TKE approaches have also been adopted for the German model, COSMO (Baldauf et al. 2011), the Japanese NHM model (Saito et al. 2006), the U.S. Navy COAMPS (Hodur 1997), and the U.S. Rapid Update Cycle model (RUC), initially using the Burk and Thompson (1989) TKE scheme.

Another trend in the 1990s among the K diffusion schemes was the increasing use of enhanced K-profile methods (similar to O’Brien 1970) along with a nonlocal or countergradient term in the vertical heat transport (similar to Deardorff 1972b). The ideas of Troen and Mahrt (1986) helped to quantify these terms as a function of surface heat flux and an ABL depth defined in terms of a bulk Richardson number. Holtslag and Boville (1993) applied this approach to replace the local scheme in the U.S. climate model, second version (CCM2), and Hong and Pan (1996) similarly replaced a local scheme in the U.S. global weather prediction model, MRF, with one that includes a countergradient term and enhanced K profile.

At the Met Office, Lock et al. (2000) also adopted countergradient term plus K-profile methods for the dry CBL and extended the boundary layer scheme to cover cloud-topped boundary layers, enhanced entrainment, and so-called top-down mixing driven by cloud-top radiative cooling. This marked the beginning of a trend to include shallow clouds in ABL schemes.

In the 2000s, the eddy-diffusion mass-flux (EDMF) approach started to be used in NWP models. This method combines local methods for vertical diffusion, either TKE based or diagnostic-K based, with a mass-flux treatment of CBL thermals using a one-dimensional plume model with a nonlocal mass flux between the surface and the thermal top (Soares et al. 2004). The ECMWF model switched from a local-K scheme to a cubic K profile, similar to those mentioned earlier with enhanced values in the middle of the ABL, plus a separate mass-flux term represented by an entraining and detraining plume transporting surface air upward (Siebesma et al. 2007). In the French models, the local vertical diffusion scheme continued to use the TKE approach, but a mass-flux plume was added (Pergaud et al. 2009), which could also handle shallow clouds. Neggers et al. (2009) also treated dry and cloud-topped thermals separately in the their EDMF scheme; this was tested by ECMWF (Köhler et al. 2011) and later adopted. Following a related but different approach, the U.S. climate model Mean Climate of the Community Atmosphere Model (CAM4) adopted Bretherton and Park’s (2009) CAM University of Washington (UW) ABL scheme, which included a diagnostic TKE and a generalized nonlocal vertical turbulence term that allows deeper moist and dry layers to mix. The CAM4 physics also considers shallow clouds separately (Park and Bretherton 2009) and includes top-down mixing.

Since 2000, the YSU ABL scheme (Hong et al. 2006), commonly used in WRF simulations, added explicit entrainment to the MRF ABL countergradient scheme modifying the K profile to stop at the top of the neutral layer and adding a term proportional to surface heat flux in an entrainment layer above. Pleim (2007) extended the Zhang and Anthes (1982) nonlocal upward transport to add a local downward flux in a so-called asymmetric convective model. Meanwhile NCEP’s rapid refresh models [RAP, High-Resolution Rapid Refresh (HRRR)] switched to another Mellor–Yamada-based TKE scheme with more sophisticated length scales (Nakanishi and Niino 2006).
Since 2010, in more recent developments the YSU ABL scheme added a top-down enhanced mixing profile driven by radiative cooling at fog (or cloud) top (Wilson and Fovell 2018), the U.S. Rapid Refresh model has been enhancing their Nakanishi–Niino TKE scheme to include an EDMF nonlocal part and shallow convection, and the GFS (formerly MRF) ABL scheme has moved from a purely countergradient nonlocal term to a hybrid countergradient and EDMF scheme based on the amount of instability (Han et al. 2016).

At present, mesoscale numerical models at fine resolution (horizontal grid spacing on the order of a few kilometers), especially if using ABL schemes with local diffusivity alone, produce enticingly realistic but only partially resolved rolls and cells to make up for the “lost” transport (e.g., Ching et al. 2014; Zhou et al. 2014). Properly representing boundary layer processes in a mesh with subkilometer-scale horizontal spacing [the PBL–LES “grey zone” or Wyngaard’s (2004) “Terra Incognita”] remains an active area of research, and several investigators are using LES (e.g., Honnert et al. 2011; Beare 2014; Shin and Dudhia 2016) or refined grids in mesoscale models (Zhou et al. 2017) to address this problem.

In conclusion, there has been a general trend from very simple dry local mixing approaches in earlier models to ABL schemes that represent nonlocal transport by thermals, cloud effects, top-down mixing, etc., thus including the full spectrum of ABL types within a single parameterization scheme and enabling improvements in representing the evolution of low cloud cover. The idea that ABL schemes should be unified with shallow cumulus schemes has been gaining popularity together with the concept of top-down radiatively driven mixing influencing both the boundary layer growth and shallow cloud development.

11. Discussion and synthesis

a. Historic synthesis

One hundred years ago, many scientists recognized that the winds in the boundary layer could be represented by some form of the Ekman spiral; that the virtual potential temperature $\theta_v$ was well mixed during the day, and that radiative cooling at night led to a deepening temperature inversion. Surface maps were becoming routine, as were kite soundings. As measuring techniques improved to allow sampling of turbulence quantities, we were able to quantify vertical transport of heat and momentum so that, by midcentury, Monin–Obukhov similarity theory provided a framework under which we could relate those fluxes to their corresponding mean profiles.

In the 1970s, ABL knowledge expanded rapidly as research emphasis shifted from the surface layer to the entire CBL, thanks to the emergence of LES, laboratory studies, and the ability to measure water vapor fluxes as well as heat and momentum fluxes from aircraft. Satellite images of cumulus clouds, tower measurements, and boundary layer clear-air radar reflectivity patterns revealed CBL rolls and cells, which were replicated in LES and explained in theoretical studies. The development of numerical models of weather and climate prompted scientists to design ambitious field studies to systematically understand the exchanges of energy and moisture at the surface and through the boundary layer under a variety of conditions. These early studies included ATEX (February 1969), BOMEX (June–July 1969), GATE (June–September 1974), and AMTEX (February 1976). We drew extensively from classical turbulence and fluid dynamics to interpret the data from these field campaigns, LES, and laboratory studies. The early boundary layer parameterizations, first driven by observations and later increasingly by LES, drew in varying degrees from “higher order” modeling, using the classical Friedmann–Keller equations (see section 2b).

The 1980s and 1990s saw increased emphasis on interaction of the boundary layer with the surface and with clouds. Seasat, launched in 1978, enabled surface-stress and wind estimates over the ocean. Progress in understanding and simulating land surface processes was greatly enhanced by documentation of soil and vegetation characteristics, much of which was satellite based. Comparison of different land surface schemes soon followed under the auspices of GEWEX as part of the PILPS (1992). Understanding of the roughness sublayer associated with canopies, ocean waves, and cities was starting to be enhanced by LES and DNS. Much of the landscape today consists of crops and urban development. Thus, it is not surprising that crop-specific surface characteristics are available as input for land surface models, or that development urban-specific and crop-specific land surface models received increased emphasis.

Cloud processes were introduced into LES in the mid-1970s; with LES of observed case studies by the early 1980s supplementing knowledge gained in aircraft observations over the tropical oceans dating back to the 1940s; and improved aircraft instrumentation enabled estimates of entrainment into stratus-topped boundary layers. While the stratus-to-cumulus transition in the tropics had been studied as far back as 1950, this became an intense area of research in the 1990s. Through GEWEX, international LES and single-column-model comparisons were started to figure out what was needed to accurately replicate observations of different types of cloud-topped boundary layers, with the goal of improving LES to the point that they could be reliable benchmarks.
in the development of parameterization schemes in weather and climate models. These efforts became more successful with time, thanks to careful case selection, improvements in LES, SCMs, and measurement technology, which, by the 1990s and 2000s, included cloud radars, Doppler and aerosol lidars, and ground-based radiometers, acoustic sounders, and Doppler radars.

The 1990s and 2000s saw significant progress in sampling the surface layer over the ocean, aided by field campaigns and the ability to measure fluxes from autonomous buoys. The impact of waves on fluxes, the subject of research in the 1950s and earlier, was being more directly addressed, both in the field (e.g., SWADE, HEXOS), and ocean-specific surface-layer models were being developed to estimate fluxes, drawing from TOGA COARE and other field programs. Flux-profile relationships were extended to increasingly strong winds, based on measurements, laboratory results, and LES. Faced with considerable scatter in experimental results in high-wind regimes, researchers focused more on the study of wave fields not in equilibrium with the wind, such as occurs in tropical cyclones.

Studies of the stable boundary layer lagged far behind, owing to difficulties in observation and LES in addition to its inherent complexity. It can be argued that there was early progress: the famous Leipzig wind profile was sampled on a weakly stable day; however, modern progress perhaps began with observations in the Arctic of a weakly stable boundary layer and its subsequent simulation. In parallel, LES at higher resolution and with more sophisticated subfilter-scale parameterizations are becoming capable of simulating more stable stratification. The SBL was the subject of a number of field programs starting in the 1990s; an Arctic program [the 1994 Beaufort Arctic Stratus Experiment (BASE)] provided the basis for an international LES comparison (Beare et al. 2006). However, these approaches continue to be limited by intermittency and even tiny surface nonuniformities.

Like the SBL, progress in understanding the evolving boundary layer has come much later. Our understanding is partially hampered by difficulties in sampling evolving boundary layers, and partially by the fact that our conception of the boundary layer has been driven by steady-state idealizations. Continued development of ABL parameterizations, elements of which are based on steady-state observations or LES, needs to be informed by comparison to observed or simulated evolving boundary layers.

b. Some lessons learned: Parameter space

1) CBL DEPTH

One kilometer is a good order-of-magnitude estimate for the depth of the CBL, but this mostly applies to the eastern United States and much of Europe in the summertime. Indeed, the daytime CBL can vary by nearly a factor of 10 in both directions, as illustrated a sampling of the literature in Fig. 9-35. Not surprisingly,
given strong surface heating and very dry conditions (hence few or no surface-based clouds), desert locations have the deepest boundary layers. Conversely, with weak surface heating, the Arctic CBL is shallow, its turbulence typically driven by shear near the surface and cloud-top cooling, and it is thus frequently transformed into two layers by decoupling of the cloud layer from the surface. In the middle are CBLs that are most often sampled and studied, which reach depths of ~1 km, with clouds in fair weather possibly extending a few kilometers more. Somewhat surprisingly, the composite mixed-layer top (not shown) reached 2 km and more over the boreal forest during BOREAS (May–August) at about 55°N (Barr and Betts 1997). This results from a combination of the length of day compensating for the large zenith angle to make total insolation not that much lower than the tropics (List 1963, Table 134) a low albedo, (Betts and Ball 1997), and high Bowen ratios (Barr and Betts 1997).

2) ENTRAINMENT (SECTIONS 4 AND 6)

Our survey of the literature reveals strong departures of the buoyancy entrainment ratio \( A_B \) from 0.2 in the early morning, when there is strong wind (or shear near CBL top); late afternoon, when surface buoyancy flux weakens and processes local to the CBL top become important; and during times of rapid change, when simply averaged negative buoyancy flux is smeared vertically. What is not always appreciated is the enormous variability of the entrainment ratios for temperature and mixing ratio even when \( A_B = 0.2 \) as illustrated in Fig. 9-36, for situations when the “negative buoyancy flux area” is sufficiently well behaved that \( A_B \) based on single-height buoyancy fluxes has meaning. From the figure, \( -A_B = (\overline{w'\theta'})_0/(\overline{w'\theta'})_b \) has been observed to have values as low as ~1.5, a reflection of important role of water vapor fluxes in moist CBLs like those over the tropical ocean. Similarly, large \( -A_q = (\overline{w'q'})_0/(\overline{w'q'})_b \) values have been observed both over land and the tropical ocean; and negative values—indicating moister air being entrained into the CBL—have been observed in the Arctic (section 9), and over the central United States when southerlies are transporting moist air from the Gulf of Mexico.

c. Some lessons learned: Physical models

Just as limited knowledge of the ABL parameter space can be a hindrance, blind use of physical models can lead to trouble. For example, flux-profile relationships rely on parameters inspired by molecular theory, where fluxes are related to local gradients: the eddy-exchange coefficients \( \left(K_M, K_H, \text{etc.}\right) \), and the turbulent Prandtl number \( \text{Pr}_T \). Yet nonlocal fluxes have been documented in both the CBL (large eddies) and SBL
(gravity waves, the eddies in Regime 2 of Sun et al. 2012). The $K$ problem is propagated further in application of the higher-order equations, which, despite some disillusionment in the 1980s, are still being used today. In CBL parameterizations, use of $K$ and higher-order models has been supplemented by, for example, adding a plume model. The picture is muddied in the case of the turbulent Prandtl number $Pr_T = K_M/K_H$, the ratio of two ill-defined functions. In the SBL (including the inversion at the top of the CBL), inclusion of gravity waves can change the trend of $Pr_T$ with stability. Similarly, several shortcomings of MOST have been enumerated, some also traced to nonlocal fluxes. Even the von Kármán constant has been the subject of vigorous debate. We can be reassured by the fact that all of these concepts have been successful to some degree; and that, by treating them with appropriate caution and respect, future researchers can develop improvements or new approaches altogether.

d. Some lessons learned: Common misconceptions or sources of confusion

Real boundary layers usually differ from the simple, quasi-steady-state boundary layers we often document or simulate. Direct heating and cooling of the air by radiation and the impact of horizontal advection are too often neglected: even when such effects are small, they can be important from a climate perspective if they are common. Likewise, CBL profiles of mean quantities often depart from their idealized forms in evolving boundary layers. Finally, whole bodies of research can be lost when phenomena are renamed—sometimes more than once, a significant problem now that researchers use key words in Internet literature searches. For example, large CBL eddies associated with strong winds have been called “longitudinal vortex rolls” (Kaimal et al. 1972), “roll vortices” (section 4), “horizontal convective rolls (HCRs)” (e.g., Weckwerth et al. 1997), and a subclass of the “turbulent organized structures (TOS)” (Kanda et al. 2004).

e. Lack of awareness and rediscovery

Lack of awareness leads to error and “firsts” that were not. Khrgian (1959, p. 226) describes two scientists in the late 1800s who, knowing from manned balloon flights that the air got colder higher up but who were unaware of adiabatic compression (discovered in the early 1800s), thought that cold air associated with winter anticyclones originated from the upper atmosphere! In the process of researching this paper, we discovered two other researchers who came close to independently deriving MOST, and one of the authors had heard of a third.30 A more current example is the rediscovery of the importance of water vapor to buoyancy. Even though known early—Espy wrote of this in 1841 and McAdie in 1915—its effects were ignored during the development of MOST and in the Businger et al. (1971) Kansas experiment paper, likely because the authors thought the effects were small. However, the same omission was made in some of the papers cited in Stull’s (1976) review paper on the entrainment ratio, where it is likely not negligible. However, measurements in AMTEX (1974) and GATE (1974) were reminders of its importance, and Brutsaert included the effects of water vapor in his definition of $L$ in his 1982 text.

f. Challenges and opportunities for the future

1) Representing surface fluxes

For the last 50-60 years, MOST has formed the basis for study of the lowest 5%-15% of the ABL; indeed, the definition of the surface layer is intimately tied to MOST. Yet shortcomings have emerged, particularly in the presence of nonlocal fluxes, in the presence of a canopy or built environment, and in the presence of ocean waves, particularly wave fields not in equilibrium with the wind (sections 3 and 5).

2) Large eddies

Above the surface layer, a significant amount of vertical transport is by eddies spanning the CBL. These eddies, with horizontal scales from 1.5 to several times the depth of the CBL, are due to flow instability, surface horizontal heterogeneity, or interaction of the CBL with tropospheric gravity waves (sections 4, 7, and 8). Their presence could make estimates of the surface-energy budget difficult. Large eddies are both difficult to document in observations, and to represent in NWP and climate models because they are only partially resolved. What representations will work for what applications?

3) Clouds

Theory, numerical modeling, and observations in concert have pushed forward our knowledge of the linkages between the boundary layer and low clouds. Understanding and representing the interplay between microphysics, turbulence, and dynamics involved in weakly to strongly coupled cloud-boundary layer systems remains a challenge (sections 8 and 9).

---

30 This likely resulted from the interruption of scientific communication during World War II.
4) NUMERICAL MODELING

Discovery, description, and modeling of boundary layer processes has benefitted from tools that we have barely touched on. From the modeling point of view, LES is benefitting from larger domains, finer resolution, the inclusion of more refined radiation and cloud-microphysics schemes, and subfilter parameterization schemes. Very high resolution LES and DNS are enhancing our understanding of near-surface and entrainment processes, and instrumental in studies of finescale processes such as airflow around cities or entrainment. Adaptive-grid methodologies (e.g., van Hooft et al. 2018) are being developed that will increase the flexibility and efficiency in studies of entrainment and evolving boundary layers. Use of immersed-boundary methods and inclusion of land surface models in LES is advancing studies of the interaction with heterogeneous surfaces. In addition, LES are now nested in mesoscale numerical models, which provides a ready-made environment for studies of surface–atmosphere interactions. These numerical approaches can enable deeper understanding of the evolving boundary layer, and the role of synoptic-scale processes, clouds, and radiation. However, LES nested in mesoscale models is in its infancy, and verification of leading-edge simulations with observations and laboratory work is essential.

5) OBSERVATIONS

Improvements in remote sensing of air motion, temperature, water vapor, trace gases, aerosols, and clouds have the potential to better sample turbulence structure and evaluate fluxes of momentum and scalars and thus increase understanding of evolving clear-air and cloud-topped boundary layers. Likewise, in situ sensing systems for measuring fluxes (including radiation), cloud particles, and aerosols, as well as new platforms like unattended aerial vehicles (UAVs), provide detail at smaller scales and at heights not always accessible by aircraft or remote sensing devices.

Older systems that have sometimes fallen into disuse are seeing new life. For example, in recent years scintillometers using temperature and water vapor–sensitive wavelengths can be combined to estimate water vapor (Meijninger et al. 2002a) and heat (Meijninger et al. 2002b) fluxes with reasonable accuracy. Similarly, data from older devices such as sodars can be used in new ways, such as interpreting profile data collected from collocated tethered or nontethered balloon platforms (e.g., Anderson 2003).

These platforms can be complemented by distributed surface sites sampling immediately above the surface, through the surface, and through the layers beneath the surface, sometimes using wireless networks of high density to sample microstructure structure in the atmosphere, the atmosphere–soil interface, and the soils. Supplementary observations and reanalysis products provide the context for a complete description of the ABL. Boundary layer scientists should look for existing observational networks, and opportunities to collaborate with citizen scientists, which often provides easier access to observational sites and maintaining and protecting instruments as well as providing extra observations. Finally, increased access of archived datasets dating back decades through online platforms offers an opportunity to test new ideas for multiple sites.

6) LABORATORY

Laboratory work has had an important role in advancing studies of dispersion and growing, entraining boundary layers. As noted in section 4, the mixed results of laboratory entrainment results are traced to a complex parameter space; such work points to the need of much larger (swimming-pool-size convection) tanks to address the entrainment problem in greater detail (Jonker and Jiménez 2014).

Acknowledgments. In a paper like this, the credit has to be shared. Thus we recognize those who drafted the individual sections. Wayne Angevine wrote the introduction (section 1) with input from Margaret LeMone. Evgeni Fedorovich wrote section 2b on early work on turbulence and fluid dynamics, with technical assistance from Elizabeth Smith; while LeMone wrote section 2a, and, with input from three informal reviewers, Peter Sullivan, Fedorovich, and Branko Kosović, section 2c on LES. LeMone and Jielun Sun wrote section 3a on the surface energy budget, LeMone wrote section 3b on the surface layer, Edward Patton wrote section 3c on canopies, and Kristina Katsaros wrote section 3d on sea–air exchanges. LeMone wrote section 4 on the CBL, Larry Mahrt wrote section 5 on the SBL, Wayne Angevine wrote section 6 on the diurnal cycle, LeMone wrote section 7 on the impact of horizontal heterogeneity, Christopher Bretherton and Don Lenschow wrote section 8 on cloud-topped boundary layers, and Michael Tjemström wrote section 9 on the boundary layer over the Arctic Ocean. For the section on representing the boundary layer in models, Jeffrey Weil wrote section 10a on dispersion, Fei Chen discussed land surface models used in weather and climate models in section 10b, and Jimy Duddia discussed ABL parameterization schemes for weather and climate models in section 10c. Wayne Angevine worked with LeMone on section 11, which is based on impressions from the earlier sections.
and the cited literature. LeMone and Angevine merged the sections into a single document; some material was moved from one section or another, and some was added; and many of the other authors reviewed other sections. In addition, Peter Sullivan provided material for the ocean discussion and Chris Fairall reviewed it. Sullivan also provided material for the LES section. Joseph Alftieri and Larry Mahrt offered advice on the horizontal heterogeneity section. Ola Persson provided useful input to the Arctic discussion. Jim Fleming supplied us with some references from around 1900 describing early work on the boundary layer and provided us with a copy of his recent book, Inventing Atmospheric Science (MIT Press, 2016), which provided useful context for us. Finally we thank Sukanta Basu and the three other reviewers for useful input, which enabled broadening and deepening the discussion in useful ways. Dudhia, Sun, and Patton are supported by NCAR, which is sponsored by the National Science Foundation; LeMone, Lenschow, and Weil are funded by TIAA/CREF with office support from NCAR; Katsaros is funded by TIAA/CREF, Fedorovich is funded by the University of Oklahoma; Tjernström is funded by the University of Stockholm; Angevine is funded by NOAA ESRL CSD; Bretherton is funded by NSF Grant AGS-1660604; Chen is funded by the NCAR Water System, USDA NIFA Grants 2015-67003-23508 and 2015-67003-23460, NSF Grant 1739705, and NOAA OAR Grant NA18OAR4590381; and Mahrt is funded by NSF Grant AGS-1614345.

REFERENCES


— —, K. J. Claffey, R. E. Jordan, C. W. Fairall, P. S. Guest, P. O. G. Pernson, and A. A. Grachev, 2000: Evaluations of the von Kármán...


Crawford, K. C., and H. R. Hudson, 1970: Behavior of winds in the...

Das, S. K., and P. A. Durbin, 2005: A Lagrangian stochastic model...

Deacon, E. L., 1953: Katabatic winds and the prevention of frost...

Crawford, L., and J. Cuxart, 2006: On the turbulence in the upper...

Conzemius, R. J., and E. Fedorovich, 2006a: Dynamics of sheared...

Cornford, C. E., 1938: Katabatic winds and the prevention of frost...

Conangla, L., and J. Cuxart, 2006: On the turbulence in the upper...

Conzemius, R. J., and E. Fedorovich, 2006b: Dynamics of sheared...

Cuxart, J., 2008: Nocturnal basin low-level jets: An integrated...


——, 2001: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.


Lu, H., and F. Porté-Agel, 2010: A modulated gradient model for
Ludwig, F. L., F. K. Chow, and R. L. Street, 2009: Effect of tur-
——, M. F. Hibberd, and M. S. Borgas, 2000: A skewed meandering
Lumley, J. L., and H. A. Panofsky, 1964: Fluctuations in inhomogeneous turbulence in the convective
Mahrt, L., 1996: The bulk aerodynamic formulation over hetero-
Manabe, S., 1969: Climate and the ocean circulation: I. The at-
McAdie, A., 1915: The theory and practise of frost fighting.

McCumber, M.C. and R. A. Pielke, 1981: Simulation of the effects

Maxwell, R. M., and N. L. Miller, 2005: Development of a coupled

——, ——, S. Zilitinkevich, I. Esau, L. Enger, and B. Grisogono,

Mauritsen, T., and G. Svensson, 2007: Observations of stably

Mauder, M., R. L. Desjardins, and J. I. MacPherson, 2008: Creating

Mason, P. J., 1988: The formation of aerially-averaged roughness


Meijninger, W. M. L., A. E. Green, O. K. Hartogensis, W. Kohsiek,

Mon. Wea. Rev.

Mellite, O. M., and S. P. Palm, 2012: Estimating te orientation and

Mellado, J. P., 2017: Cloud-top entrainment in stratocumulus to

Mellor, G., and T. Yamada, 1974: A hierarchy of turbulence
closure models for planetary boundary layer. J. Atmos. Sci.,

Melfi, S. H., and S. P. Palm, 2012: Estimating the orientation and

McGavin, R. E., and M. J. Vetter, 1963: Radio refractometry and

McGillivray, B. and Coauthors, 2011: Indirect and Semi-


May, A., J. H. Yen, and T. S. Saha, 2013: Understanding large-


McCambridge, M. C., and R. A. Pielke, 1981: Simulation of the effects

McCumber, M.C. and R. A. Pielke, 1981: Simulation of the effects

McFarquhar, G. M., and Coauthors, 2011: Indirect and Semi-


McGibbon, J., and C. S. Bretherton, 2017: Skill of ship-following

McGillivray, B. and Coauthors, 2011: Indirect and Semi-


Meehl, G. A., 1997: Cloud-top entrainment in stratocumulus to

Melting, R., 1979: An empirical parameterization of surface

Meehl, G. A., 1997: Cloud-top entrainment in stratocumulus to

Melting, R., 1979: An empirical parameterization of surface

Meehl, G. A., 1997: Cloud-top entrainment in stratocumulus to

Melting, R., 1979: An empirical parameterization of surface

Meehl, G. A., 1997: Cloud-top entrainment in stratocumulus to

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface

Melting, R., 1979: An empirical parameterization of surface


