A. Collaborative Research: Droplet transport in the vicinity of breaking waves: Experiments and simulations

NSF Award Number: OCE-1829515
PI: David Richter, University of Notre Dame
Submission date: March 23, 2020
Total request: 10 million CPU-hours, 10TB Campaign Storage

B. Overview of Project

The NSF-funded project associated with this request is broadly focused on better understanding airborne sea spray droplet transport immediately after formation in the wave boundary layer. The generation of spray and aerosols at the ocean surface leads to multiple influences on Earth's weather and climate, and accurately predicting their rate of production is essential for a wide variety of applications. While the past several decades has seen great improvements in reducing the uncertainty of so-called sea spray generation functions (SSGFs) for small droplets, these estimates remain highly unconstrained for large droplets (see Figure 1 and Veron [2015]). The collaborative research project aims at characterizing and quantifying one of the biggest sources of this uncertainty: a near complete lack of knowledge regarding droplet transport in the turbulent airflow around breaking surface waves (i.e. within 1-2 meters of the ocean surface). At the air-sea interaction laboratory at the University of Delaware (UD), controlled, repeatable breaking wave events are being produced, where measurements of size-resolved droplet concentration and droplet velocity are yielding direct estimates of production fluxes. At the University of Notre Dame (UND), measured waveforms and droplet production rates are being used as inputs into large eddy simulations (LES) configured to recreate laboratory conditions. Droplet statistics will be directly compared between experiments and simulations, and the simulations will provide a means for investigating momentum and thermodynamic exchange rates, as well as upscaling to field conditions. This computational request is in support of the numerical simulations at UND, where the PI's NTLP model (NCAR Turbulence with Lagrangian Particles) will be used in its "wavy" configuration to mimic conditions in the wind-wave tank at UD, including the experimentally-measured waveforms and initial droplet size/velocity spectra. The underlying code has been routinely used at production scale using both CISL resources as well other supercomputing centers (e.g. DoD, DoE). The code is ready for immediate use, and will use the proposed computational resources (both computing and data storage) efficiently. The computational plan outlined below provides for a series of idealized wave simulations, as well as a series of experimentally-matched conditions as data is available from UD, to investigate droplet transport, lifetime, and constraints on the SSGF for large droplets.

C. Science Objectives

In the high-wind marine atmospheric boundary layer, significant amounts of sea spray can be ejected into the air; The degree to which these droplets modify near-surface turbulence and thermodynamics is largely unknown. One of the big reasons for this continued ambiguity is in our lack of understanding of certain droplet properties immediately after being produced

via wave breaking or bubble bursting. These properties include their trajectories, their lifetimes and change in temperature/size due to evaporation. Characterizing this droplet transport is no easy task, and requires understanding both the turbulent air flow around waves (a very difficult goal in its own right), as well as the evolution of droplets as they traverse this turbulent field. Only recently have experimental measurements been able to probe the turbulence structure in the immediate vicinity of waves [Reul et al., 2008, Buckley and Veron, 2016], and these have been complemented by LES [Sullivan et al., 2018, Husain et al., 2019] to better understand the consequences of wave-induced form drag, wind-wave alignment, and airflow separation (see Figure 2).

How these complex turbulent fields dictate the motion of droplets immediately after formation plays a huge role when trying to estimate bulk aerosol production (via the SSGF) or predict droplet influence on air-sea momentum, energy, moisture, and gas transfer. Without a proper understanding of how droplets and aerosols go from formation to free atmospheric transport, estimating their production and thermodynamic influence is fundamentally limited [Andreas et al., 2015, Fairall et al., 1994].

It is the objective of the overarching NSF project, therefore, to combine the PI's spray droplet/LES model with recent experimental developments in near-surface imaging using UD's wind-wave tank. In one set of simulations, idealized conditions will be prescribed in the LES in order to determine the fundamental physical processes which determine droplet transport in the wavy boundary layer. Statistics will be collected on droplet lifetimes and phase-averaged concentrations, which allow for the modification of traditional flux-profile methods for estimating surface fluxes [Lewis and Schwartz, 2004] and formulating upscaled stochastic models [Mueller and Veron, 2014]. These, experimental measurements from UD will provide boundary and initial conditions for the LES, which will in turn be used to collect important statistics on droplet lifetime and temperature/radius evolution that can be validated by the experiments. This combined information will ultimately be ultimately used to scale to field conditions (not part of this computational request) and refine existing models of bulk spray-mediated heat and moisture transport.

Therefore the specific purpose of this computational request is twofold: (1) to first provide resources for conducting an idealized suite of simulations at sufficiently high grid resolution so that the basic properties of droplet-laden wave turbulence can be accurately identified, and (2) to provide resources which will allow for combining the LES simulations with the experimental measurements. Since resolving near-surface turbulence is critical for accurately capturing the Lagrangian trajectory of individual droplets, grid resolutions will be used which minimize the need for relying on sub-grid droplet transport, requiring grid spacing on the order of sub-centimeter near the wavy surface.

D. Computational Experiments and Resource Requirements

Computational Experiments

1. Numerical Approach

The numerical approach for the proposed computations is LES coupled with Lagrangian tracking for the individual droplets. As noted above, the code is built upon the NCAR LES model [Sullivan and Patton, 2011, Moeng, 1984], and has been extended to track the trajectories of up to hundreds of millions of individual droplets [Peng and Richter, 2017,

Sweet et al., 2018]. Briefly, the flow component of the code integrates the incompressible Navier-Stokes equations for mass, momentum, and energy conservation. In the "wavy" configuration, a curvilinear coordinate transformation is used to map the physical grid to a Cartesian computational grid (see Figure 3(a)). The details of this curvilinear coordinate transformation and its implementation in the NCAR LES code can be found in Sullivan et al. [2018] and Sullivan et al. [2014].

The droplet solver tracks individual particles according to mass, momentum, and energy conservation as well, and these droplets are independent of the flow computational grid (see Figure 3(b)). Each particle's location $\vec{x}_p$, velocity $\vec{v}_p$, temperature T_p, and mass m_p is determined by the conservation equations. A droplet's velocity is determined by gravity and the surrounding flow velocity, the temperature is determined by the evaporation rate and the local air temperature, and the droplet mass is determined by the local relative humidity [Helgans and Richter, 2016]. Any gain/loss of momentum, vapor mass, or energy is conserved between the droplet and air phases, reflecting a two-way coupling between air and spray.

At the lower boundary, the wave shape is prescribed by the user, and can take an idealized, functional form (e.g., monochromatic sinusoidal) or purely empirical (e.g., $h(x, y, t)$ data from the experiments where h is the water surface height). The curvilinear coordinate transformation is dynamic in time, so that waveforms can move along the bottom with any desired dispersion relation and wave orbital velocities. For the first set of simulations, idealized, linear waves will be used to survey the basic structure of turbulence and droplet transport past simplified waveforms (see Figure 3(a)), and the second set of simulations will use measured waveforms from the UD wind-wave tank (see Figure 4). Droplets can be initialized at the bottom surface at any desired location, with a set size and initial velocity. The first set of simulations will make assumptions regarding the initial droplet conditions, while the second set of simulations will use experimentally measured values.

The code is written in Fortran with MPI using a two-dimensional domain decomposition in the Cartesian x, y, and z directions. The decomposition for the particle solver is done differently than that for the flow solver to maintain load balancing (i.e., uniform number of particles on each processor). Code I/O is done efficiently using parallel write/read statements, and takes up negligible amounts of time in the course of a typical simulation. The code will be run primarily by a graduate student of the PI from the University of Notre Dame under the direct guidance of the PI.

2. Computational Experiments

As part of the initial stages of the NSF project, the PI and his group have recently conducted a preliminary study using simulations similar to those proposed. As detailed in Richter et al. [2019], the spatial distribution of spray droplets, as well as Lagrangian statistics including lifetime, maximum height achieved, and water re-entry position, are all a strong function of wave parameters (wave age and steepness in particular) as well as droplet size. This study was intended to be a first step in the larger NSF project, and sets the stage for the proposed large-scale computations which make up the current request.

As such, the computational grid used in Richter et al. [2019] was limited to $[N_x, N_y, N_z] = [128, 128, 128]$, which for our domain size translates to roughly $\mathcal{O}(1\,\mathrm{cm})$ in the cross-channel and vertical directions, and nearly $\mathcal{O}(10\,\mathrm{cm})$ in the streamwise direction. The total droplet numbers were likewise limited to $\mathcal{O}(10^6)$. Since our goal, however, is to compare to the UD

wind-wave tunnel experiments, significantly higher resolution and domain extent is needed, particularly in the streamwise direction, where many more grid points are required in order to resolve a wider (i.e. more realistic) spectrum of wave modes at the lower surface. Similarly, a larger number of droplets are desired in order to improve statistical convergence and allow for a more realistic droplet size spectrum.

Therefore in order to more faithfully capture the dynamics in the UD wind-wave tunnel, the proposed simulations require grid spacings of $[N_x, N_y, N_z] = [2048, 256, 256]$ and particle numbers of $N_p = 10^8$. The streamwise extent is determined by the experimental measurement window and typical wavelengths and wave spectra seen in past experiments. The particle number is taken as an upper bound for what should be expected at the wind tunnel speeds.

The proposed cases are outlined below, *but here we emphasize that the computational cost of each simulation is dictated entirely by the number of grid points and droplets, and not by the setup.* In other words, the computational expense will not change between our "idealized" setup versus the "tank" setup.

The proposed simulations can be broken into two broad categories, each with a certain set of parameters to be varied:

1. **Idealized simulations, similar in nature to those in Richter et al. [2019], but with both a larger domain and higher resolution**. Here, two key wind/wave parameters will be varied: wave age and wave steepness. Wave ages of $c/u_* = 1, 5, 10$ will be used, each with wave slopes of $ak = 0.1, 0.3, 0.5$, where wave age is a measure of how fast the winds are (u_* is the friction velocity) relative to the wave speed (c is the wave phase speed) and wave slope is a measure of the steepness of the wave (a is the wave amplitude and k is the wavenumber). This, plus a "flat" reference case (no lower wave) results in 10 total idealized simulations.

2. **UD wave tank simulations, designed to mimic conditions in the tank**. Here, there are two separate experiments being conducted, each varying wind speed as the primary variable:

   • Wind-forced waves, where the water surface is exposed to winds ranging from $3-15\,\mathrm{m/s}$. The waves which naturally develop will be measured in the laboratory and their shape will be provided to the simulations as a lower boundary condition. We request 5 simulations at 5 different wind speeds in this range[1].

   • Mechanically forced waves, where waves are produced by an upstream paddle and exposed to winds in the same $3-15\,\mathrm{m/s}$ wind speed range. These can be designed to break at a specific point in the tunnel, and be used to vary wave age beyond what can be examined with wind-forced waves. As above, we request 5 simulations at 5 different wind speeds for the mechanically forced waves.

_____

[1]We admit that the number 5 is somewhat arbitrary; however, we anticipate that the wave field and spray distribution will change rapidly over these wind speeds. Furthermore, since the experiments are only just beginning, it has not yet been determined which wind speeds will be used, and 5 provides sufficient flexibility if/when changes are made in the experiments

In total, this amounts to 20 simulations, each with the same grid size and roughly the same number of droplets (the exact droplet numbers will depend on the measured droplet production rates for the wave tank simulations). Each simulation will be run beyond statistical equilibrium (a total of roughly 15 minutes in simulation time), necessitating 100,000 time steps. In each simulation, statistics will be collected on droplet concentrations, lifetimes, and thermodynamic exchange rates. Models for spray transport and air-sea exchange will then be tested directly against the highly-resolved simulations. For the wave tank simulations, validations will be performed on local computational resources before the production simulations proposed above.

3. Code Performance

The NTLP code has been routinely run on CISL resources, including past projects on Yellowstone and Cheyenne, and has repeatedly demonstrated efficiency and scalability. Using a recent small university allocation, timing tests have been performed on Cheyenne up to 4608 cores. Figure 5 shows that both the flow solver and the Lagrangian droplet solver scale well on the Cheyenne system. Furthermore, these tests indicate that the timing amounts to:

- 1.3×10^{-11} seconds of wall clock time per time step per Cheyenne core per grid point

- 4.3×10^{-12} seconds of wall clock time per time step per Cheyenne core per droplet.

Resource Requirements

1. HPC

From the numbers above, the 20 proposed simulations with $[N_x, N_y, N_z] = [2048, 256, 256]$ and $N_p = 10^8$ equals 10M cpu-hours. Half of these make up the idealized simulations, and half make up the wave tank simulations. We anticipate using 2304 Cheyenne cores for the production runs.

2. Scratch storage

In total, 35TB of scratch storage is requested to store raw, uncompressed output as the simulations are running. The bulk of this (30 TB) is from the full "volumes" written out at every 3000 time steps (33 volume files total, each at 45 GB), where the required flow and particle variables are written in double-precision binary format. These are necessary both for collecting statistics and other types of analysis (e.g. visualization), as well as re-starting the simulations. The remainder (5 TB) is for the compressed statistics files which will be generated as the simulations progress.

3. Campaign storage

The requested campaign storage will be no more than 10 TB. The raw output stored in scratch will be analyzed and compressed in multiple formats, totaling roughly 8 TB. In addition, two of the full raw volume files will be stored on campaign so that restarts from a fully-developed field can be easily performed if additional statistics are needed in the future. This amounts to roughly 2 TB. The rest of the raw scratch data will be discarded or moved to local storage as the analysis is completed.

4. Data analysis and visualization

Nearly all of the postprocessing is done using standard software: python, Matlab, and Fortran. The standard allocation of 10,000 cpu-hours on the DAV cluster will suffice for the various analysis steps.

E. Data Management Plan

As outlined above, the total scratch disk space required is roughly 35 TB. This allows for full, uncompressed restart volumes, planes of data for visualization, and particle information to be stored during each run. This stored data will be used to probe flow variables in instantaneous snapshots, make time animations of various flow variables, and restart the simulation when necessary. As the simulations are completed, this raw data will be analyzed and compressed in a variety of forms (depending on each type of analysis). Once each simulation is complete, the postprocessed, compressed results will be stored in campaign storage and the original raw volumes will be discarded within the scratch purge window. Only 2 full restart volumes will be retained on campaign, in the event that additional simulation time is needed in the future.

According to the data management plan contained in the original NSF proposal, the raw simulation data will only be made available under special circumstances due to its specialized format and the size of the individual files. Instead, the PI has and will continue to make the code and all necessary initial condition files and instructions available to the scientific community. Postprocessed data, particularly that data which is published, will then be stored on Notre Dame's new CurateND platform — a digital database which specializes in making data and other digital work publicly available.

F. Accomplishment Report

The PI has extensive experience running MPI-based codes on HPC resources. Among many smaller allocations and past projects, the PI was the sole user of a 5.2 million core-hour allocation granted as part of the Accelerated Scientific Discovery at the initiation of the Yellowstone machine in the fall of 2012. The simulations performed during the ASD allocation provided a unique dataset which directly led to a series of publications [Richter and Sullivan, 2014b, Richter, 2015, Richter and Sullivan, 2014a] concerning the upscale effect of small droplets on the turbulent energy cascade. In addition, the PI was the lead on a previous University Large Allocation (UNDM0004) starting in 2016 (on a separate NSF project from the present request), which successfully used 2.5 million core-hours to study sensible and latent heat exchange using the NTLP code. This also resulted in a series of publications which have helped step forward our understanding of air-sea transfer with spray [Peng and Richter, 2017, 2019, Under review].

G. References

E. L. Andreas, L. Mahrt, and D. Vickers. An improved bulk air-sea surface flux algorithm, including spray-mediated transfer. *Quarterly Journal of the Royal Meteorological Society*, 141:642–654, 2015. doi: 10.1002/qj.2424.

M. P. Buckley and F. Veron. Structure of the airflow above surface waves. *Journal of Physical Oceanography*, 46:1377–1397, 2016. ISSN 0022-3670. doi: 10.1175/JPO-D-15-0135.1.

C. W. Fairall, J. D. Kepert, and G. J. Holland. The effect of sea spray on surface energy transports over the ocean. *The Global Atmosphere and Ocean System*, 2:121–142, 1994.

B. Helgans and D. H. Richter. Turbulent latent and sensible heat flux in the presence of evaporative droplets. *International Journal of Multiphase Flow*, 78:1–11, 2016. doi: 10.1016/j.ijmultiphaseflow.2015.09.010.

N. T. Husain, T. Hara, M. P. Buckley, K. Yousefi, F. Veron, and P. P. Sullivan. Boundary layer turbulence over surface waves in a strongly forced condition: LES and observation. *Journal of Physical Oceanography*, 49:1997–2015, 2019. doi: 10.1175/JPO-D-19-0070.1.

E. R. Lewis and S. E. Schwartz. *Sea Salt Aerosol Production: Mechanisms, Methods, Measurements, and Models - a Critical Review.* American Geophysical Union, 2004.

C.-H. Moeng. A large-eddy-simulation model for the study of planetary boundary-layer turbulence. *Journal of the Atmospheric Sciences*, 41(13):2052–2062, 1984. doi: 10.1175/1520-0469(1984)041¡2052:ALESMF¿2.0.CO;2.

J. A. Mueller and F. Veron. Imact of sea spray on air-sea fluxes. Part I: Results from stochastic simulations of sea spray drops over the ocean. *Journal of Physical Oceanography*, 44:2817–2834, 2014. doi: 10.1175/JPO-D-13-0245.1.

T. Peng and D. Richter. Influence of evaporating droplets in the turbulent marine atmospheric boundary layer. *Boundary-Layer Meteorology*, 165:1–22, 2017. doi: 10.1007/s10546-017-0285-7.

T. Peng and D. Richter. Influences of poly-disperse sea spray size distributions on model predictions of air-sea heat fluxes. *Journal of Geophysical Research Atmospheres*, Under review.

T. Peng and D. H. Richter. Sea spray and its feedback effects: Assessing bulk algorithms of air-sea heat fluxes via direct numerical simulations. *Journal of Physical Oceanography*, 49:1403–1421, 2019. doi: 10.1175/JPO-D-18-0193.1.

N. Reul, H. Branger, and J. P. Giovanangeli. Air flow structure over short-gravity breaking water waves. *Boundary-Layer Meteorology*, 126:477–505, 2008. doi: 10.1007/s10546-007-9240-3.

D. H. Richter. Turbulence modification by inertial particles and its influence on the spectral energy budget in planar Couette flow. *Physics of Fluids*, 27:063304, 2015. doi: 10.1063/1.4923043.

D. H. Richter and P. P. Sullivan. The sea spray contribution to sensible heat flux. *Journal of the Atmospheric Sciences*, 71(2):640–654, 2014a. doi: 10.1175/JAS-D-13-0204.1.

D. H. Richter and P. P. Sullivan. Modification of near-wall coherent structures by inertial particles. *Physics of Fluids*, 26:103304, 2014b. ISSN 1070-631. doi: 10.1063/1.4900583.

D. H. Richter, A. E. Dempsey, and P. P. Sullivan. Turbulent transport of spray droplets in the vicinity of moving surface waves. *Journal of Physical Oceanography*, 49:1789–1807, 2019. doi: 10.1175/jpo-d-19-0003.1.

P. P. Sullivan and E. G. Patton. The effect of mesh resolution on convective boundary layer statistics and structures generated by large-eddy simulation. *Journal of the Atmospheric Sciences*, 68:2395–2415, 2011. doi: 10.1175/JAS-D-10-05010.1.

P. P. Sullivan, J. C. McWilliams, and E. G. Patton. Large-eddy simulation of marine atmospheric boundary layers above a spectrum of moving waves. *Journal of the Atmospheric Sciences*, 71:4001–4027, 2014. doi: 10.1175/JAS-D-14-0095.1.

P. P. Sullivan, M. L. Banner, R. P. Morison, and W. L. Peirson. Turbulent flow over steep steady and unsteady waves under strong wind forcing. *Journal of Physical Oceanography*, 48:3–27, 2018. doi: 10.1175/JPO-D-17-0118.1.

J. Sweet, D. H. Richter, and D. Thain. GPU acceleration of Eulerian-Lagrangian particle-laden turbulent flow simulations. *International Journal of Multiphase Flow*, 99:437–445, 2018. doi: 10.1016/j.ijmultiphaseflow.2017.11.010.

F. Veron. Ocean spray. *Annual Review of Fluid Mechanics*, 47:507–538, 2015. doi: 10.1146/annurev-fluid-010814-014651.
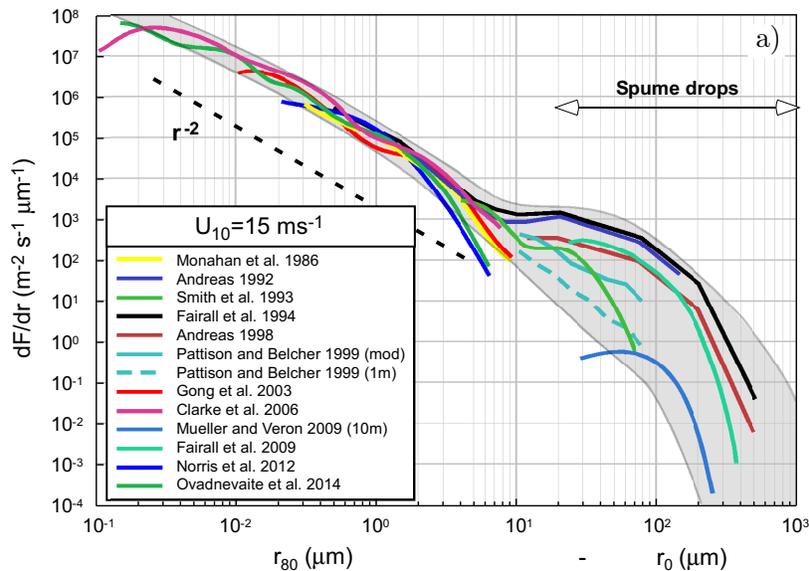
H. Figures

Figure 1: Adapted from Veron [2015]. Compilation of SSGFs at wind speeds of 15 m/s, highlighting large uncertainties in the spume droplet regime (i.e. large droplets torn from wave crests).
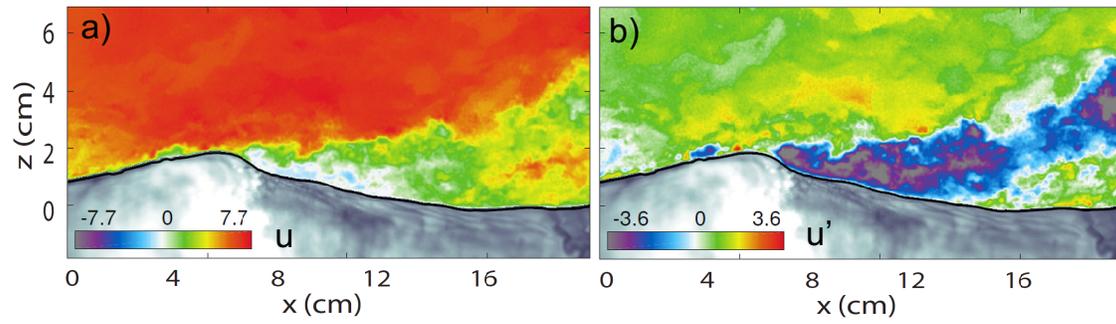
Figure 2: Adapted from Buckley and Veron [2016]. (a) Wind speed and (b) velocity fluctuation from experimental measurements of airflow over a moving surface wave.
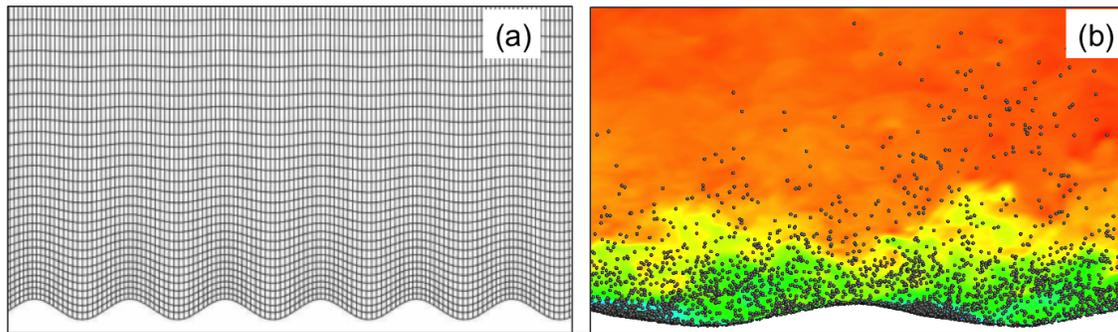

Figure 3: (a) Computational grid in physical space, which is then mapped to a Cartesian space for calculation. (b) Lagrangian droplets (dots) released from an idealized waveform at the surface.
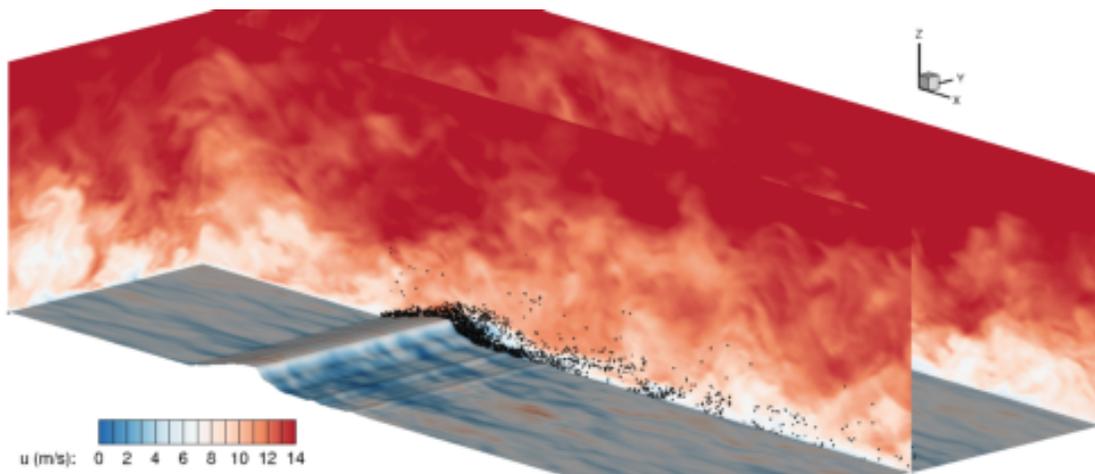

Figure 4: LES simulations with Lagrangian droplets using a measured waveform from a previous UD wave tank experiment.
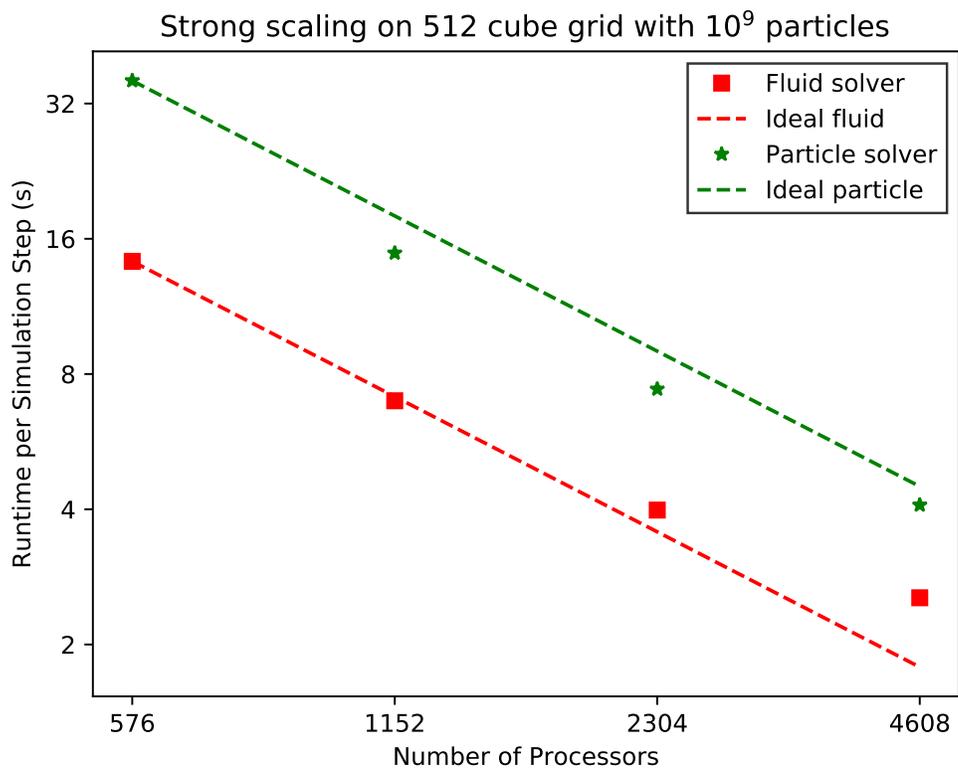
9

Figure 5: Strong scaling curve for the proposed code on Cheyenne. The particle solver and flow solver have been separated, and each independently display favorable strong scaling for a grid similar to that proposed.