# REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)
01-04-2016

2. REPORT TYPE
Final Report

3. DATES COVERED (From - To)
1-Aug-2011 - 31-Jul-2015

4. TITLE AND SUBTITLE
Final Report: Environmental Sciences:YIP:Combining Remotely Sensed Vegetation Data & Ecohydrologic Process Models to Improve Estimation of Root Zone Moisture at Spatial Scales Relevant to the Army

5a. CONTRACT NUMBER
W911NF-11-1-0310

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Alejandro N. Flores

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Boise State University
1910 University Drive

Boise, ID 83725 -1135

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
58471-EV.9

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

Soil moisture is a critical land surface hydrology and battlespace environment variable. As a hydrologic variable, it controls the partitioning of incoming energy into sensible and latent heat fluxes. In the battlespace environment soil moisture impacts the Army mobility. This report summarizes research activities, findings, and implications for a four-year effort to advance fundamental understanding of coupled land-atmosphere systems and the ability to predict the dynamics of these systems at spatiotemporal resolutions of interest to the Army using models and data. Advances were made on three fronts: (1) fusion of multi-sensor remote sensing imagery to improve characterization

15. SUBJECT TERMS
Hydrology, Ecohydrology, Land-atmosphere interaction, Data assimilation, Remote sensing

16. SECURITY CLASSIFICATION OF:
a. REPORT UU
b. ABSTRACT UU
c. THIS PAGE UU

17. LIMITATION OF ABSTRACT
UU

15. NUMBER OF PAGES

19a. NAME OF RESPONSIBLE PERSON
Alejando Flores

19b. TELEPHONE NUMBER
208-426-2903

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

## Report Title

Final Report: Environmental Sciences:YIP:Combining Remotely Sensed Vegetation Data & Ecohydrologic Process Models to Improve Estimation of Root Zone Moisture at Spatial Scales Relevant to the Army

## ABSTRACT

Soil moisture is a critical land surface hydrology and battlespace environment variable. As a hydrologic variable, it controls the partitioning of incoming energy into sensible and latent heat fluxes. In the battlespace environment soil moisture impacts the Army mobility. This report summarizes research activities, findings, and implications for a four-year effort to advance fundamental understanding of coupled land-atmosphere systems and the ability to predict the dynamics of these systems at spatiotemporal resolutions of interest to the Army using models and data. Advances were made on three fronts: (1) fusion of multi-sensor remote sensing imagery to improve characterization of terrestrial vegetation at hillslope scales (30 m), (2) use of machine learning algorithms to predict net ecosystem exchange using variables that can be remotely sensed, and (3) use of coupled land-atmosphere models to derive environmental forcings at high spatial resolutions in complex landscapes. This Young Investigator Program project engaged both undergraduate and graduate students and generated publications and datasets that have been used to leverage additional research funding. Project activities and findings also reveal new research avenues that are of potential significance for Army applications to be pursued in the future.

---

## Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

### (a) Papers published in peer-reviewed journals (N/A for none)

| Received | | Paper |
|---|---|---|
| 04/01/2016 | 7 | Patrick R. Kormos, James P. McNamara, Mark S. Seyfried, Hans Peter Marshall, Danny Marks, Alejandro N. Flores. Bedrock infiltration estimates from a catchment water storage-based modeling approach in the rain snow transition zone,<br>Journal of Hydrology, (06 2015): 0. doi: 10.1016/j.jhydrol.2015.03.032 |
| 04/01/2016 | 8 | Patrick R. Kormos, Danny Marks, James P. McNamara, H.P. Marshall, Adam Winstral, Alejandro N. Flores. Snow distribution, melt and surface water inputs to the soil in the mountain rain–snow transition zone,<br>Journal of Hydrology, (11 2014): 0. doi: 10.1016/j.jhydrol.2014.06.051 |
| 11/23/2014 | 5 | Reggie D. Walters, Katelyn A. Watson, Hans-Peter Marshall, James P. McNamara, Alejandro N. Flores. A physiographic approach to downscaling fractional snow cover data in mountainous regions,<br>Remote Sensing of Environment, (09 2014): 413. doi: 10.1016/j.rse.2014.07.001 |
| 11/23/2014 | 6 | Alejandro N. Flores, Dara Entekhabi, Rafael L. Bras. Application of a hillslope-scale soil moisture data assimilation system to military trafficability assessment,<br>Journal of Terramechanics, (02 2014): 53. doi: 10.1016/j.jterra.2013.11.004 |
| **TOTAL:** | | **4** |

**Number of Papers published in peer-reviewed journals:**

---

## (b) Papers published in non-peer-reviewed journals (N/A for none)

| Received | Paper |
|---|---|
| **TOTAL:** | 1 |

**Number of Papers published in non peer-reviewed journals:**

---

## (c) Presentations

**Number of Presentations:** 0.00

---

## Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

| Received | Paper |
|---|---|
| **TOTAL:** | |

**Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):**

---

## Peer-Reviewed Conference Proceeding publications (other than abstracts):

| Received | Paper |
|---|---|
| **TOTAL:** | |

**Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):**

---

## (d) Manuscripts

| Received | Paper |
|---|---|

**TOTAL:**

**Number of Manuscripts:**

---

## Books

| Received | Book |
|---|---|

**TOTAL:**

| Received | Book Chapter |
|---|---|

**TOTAL:**

## Patents Submitted

---

## Patents Awarded

---

## Awards

NSF CAREER Award, Idaho Business Review Accomplished Under 40 (nominated for 2016)

---

## Graduate Students

| NAME | PERCENT_SUPPORTED | Discipline |
|---|---|---|
| Katelyn Watson | 0.50 | |
| **FTE Equivalent:** | **0.50** | |
| **Total Number:** | **1** | |

## Names of Post Doctorates

| NAME | PERCENT_SUPPORTED |
|---|---|
| Qingtao Zhou | 1.00 |
| **FTE Equivalent:** | **1.00** |
| **Total Number:** | **1** |

## Names of Faculty Supported

| NAME | PERCENT_SUPPORTED | National Academy Member |
|---|---|---|
| Alejandro Flores | 0.25 | No |
| **FTE Equivalent:** | **0.25** | |
| **Total Number:** | **1** | |

## Names of Under Graduate students supported

| NAME | PERCENT_SUPPORTED |
|---|---|
| **FTE Equivalent:** | |
| **Total Number:** | |

## Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

| | |
|---|---|
| The number of undergraduates funded by this agreement who graduated during this period: ...... | 0.00 |
| The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... | 1.00 |
| The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... | 0.00 |
| Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... | 0.00 |
| Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... | 0.00 |
| The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense ...... | 0.00 |
| The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... | 0.00 |

## Names of Personnel receiving masters degrees

| NAME |
|---|
| **Total Number:** |

## Names of personnel receiving PHDs

| NAME | |
|---|---|
| Katelyn Watson | |
| **Total Number:** | **1** |

## Names of other research staff

| NAME | PERCENT_SUPPORTED |
|---|---|
| Peter Olsoy | 0.50 |
| **FTE Equivalent:** | **0.50** |
| **Total Number:** | **1** |

### Sub Contractors (DD882)

## Inventions (DD882)

## Scientific Progress

See Attachment

## Technology Transfer

# Young Investigator Program: Environmental Sciences: Combining Remotely Sensed Vegetation Data and Ecohydrologic Process Models to Improve Estimation of Root Zone Moisture at Spatial Scales Relevant to the Army

Award No. W911NF1110310
Principal Investigator: Alejandro N. Flores, Ph.D.

## Final Report Narrative

## 1. Background and Introduction

Soil moisture is a critical land surface hydrology variable, as well as an important battlespace environmental variable. As a hydrologic variable, soil moisture controls the partitioning of incoming energy into sensible and latent heat fluxes. In the battlespace environment soil moisture impacts the mobility and deployability of troops and materiel. For Army and many hydrologic applications, the scale at which soil moisture information can be obtained is of critical importance. In particular, for Army applications it is desirable to have reliable estimates of soil moisture at spatial resolutions of hillslopes (e.g., 10s to 100s m) over areas corresponding to large watersheds (e.g., 100s to 1000s km$^2$). Given the potential realm of Army operations, moreover, it is necessary to have a generalized framework for constraining soil moisture at these resolutions and scales anywhere in the world.

There are multiple data streams that can provide information that is useful in constraining soil moisture, yet there is no single data source that is capable of providing soil moisture information at the relevant spatiotemporal resolutions and scales.

Direct observations of soil moisture at these spatial resolutions and areal extents is generally not possible with current technologies. Remote sensing is an attractive option because spaceborne satellites can provide global observational coverage at regular revisit intervals. Spaceborne sensors, however, provide only indirect soil moisture information (e.g., surface reflectance or radiance) in particular wavelengths to which soil moisture is sensitive. The development of retrieval algorithms is needed to invert the observed geophysical properties for soil moisture. Developing these algorithms in a robust way requires ground-based observations of soil moisture across a range of soil types, land use/land cover types, and under a variety of climatological conditions. Current soil moisture platforms are also associated with spatial resolutions that are too coarse for typical Army applications. A notable milestone in soil moisture remote sensing is the NASA Soil Moisture Active-Passive (SMAP) mission, which was designed with microwave radar capable of resolving radar backscatter in the L-band (1.26 GHz) at spatial resolutions of 2-3 km globally. Shortly after launch, however, the radar failed leaving SMAP only with a passive radiometer to observe soil moisture at spatial resolutions of approximately 36 km.

Spatially distributed, process-based ecohydrologic models are capable of providing real-time estimates of soil moisture. Such models also serve to fuse multiple-scale, high resolution geospatial data (e.g., digital elevation models, vegetation fraction) to simulate the redistribution of input precipitation in space and time. But these models are subject to uncertainties in the input environmental forcings (e.g., precipitation, temperature, humidity, etc.), the fidelity of those parameterizations to underlying physical processes, and the associated values of parameters input to those models. Particularly problematic for model-based forecasting of soil moisture at tactical scales are the lack of globally available numerical weather forecasts at spatial resolutions that resolve heterogeneity in complex, mountain terrain (e.g., $\leqq 3$ km). As a result of these uncertainties and limitations in current availability of model inputs, predictions of soil moisture based on hydrologic models will inevitably be in error.

Data assimilation and data fusion provide mathematical frameworks through which predictions derived from models can be constrained to observations from remote sensing (as well as ground-based) platforms. Although there is a range in complexity of these various approaches, the thrust of these methods is to provide mechanisms for constraining models to observations. The overarching objective of this Young Investigator Program (YIP) award is to improve the ability to estimate surface water and energy balance fluxes through the use of models and data. Below are outlined a narrative of the activities undertaken during this grant and how they integrate (Section 2), a summary of products produced and/or in progress and the implications of these products and knowledge for Army operations (Section 3), and an overview of future potential extensions of this research for applications that are of interest to the Army.

## 2. Activities

### *2.1 Improving the spatiotemporal resolution of remotely sensed vegetation products*

Remote sensing of terrestrial vegetation provides critically important information about the dynamics of terrestrial ecosystems and, in water-limited ecosystems, indirect information about terrestrial hydrology. Multi-spectral remote sensing in the 400 to 2500 nm (visible to shortwave infrared) wavelengths is a particularly popular technique for characterizing terrestrial ecosystems and hydrologic dynamics. A number of existing and planned platforms provide multispectral data across a range of spatial and temporal resolutions. NASA’s Landsat missions have provided continued remote sensing of terrestrial ecosystems since the 1970s. Contemporary Landsat platforms are associated with a 30 m spatial resolution in 7 bands at a temporal revisit rate of 16 days. The historical temporal coverage of Landsat, moreover, provides important information from which long-term trends and changes in terrestrial ecosystems can be observed. The MODerate-resolution Imaging Spectroradiometer (MODIS) is a similar instrument that is deployed both on NASA’s Terra and Aqua satellites. MODIS provides spectral information at spatial resolutions down to 250 m with 1-2 overpasses per day. Image composites of vegetation indices like the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) are available at 8 days. Leveraging observations from both Terra and Aqua satellites yields vegetation index composites at temporal revisit rates of 4 days.

One significant downside of spectral remote sensing methods is that clouds can obscure the ability to observe terrestrial vegetation. For Army applications, moreover, both of the most commonly used remote sensing platforms (Landsat and MODIS) are associated with significant limitations. With a 16 day interval between successive observations – longer if cloud cover significantly degrades the visibility of the land surface – Landsat imagery can miss important changes in terrestrial vegetation. Key facets of vegetation dynamics that could be missed include phenological changes (green-up or senescence) and disturbance due to fire. Although MODIS provides more frequent observations in time, the spatial resolution of the associated vegetation products is too coarse for many applications

One possible way to leverage the strengths of each of these remote sensing platforms to compensate for their corresponding weaknesses would be to use data fusion approaches to combine MODIS and Landsat data to yield a combined vegetation index dataset with a spatial resolution of 30 m and a temporal resolution of 4 days. STAR-FM, an open source software framework, provides such an opportunity (Gao et al., 2006).

In the Dry Creek Experimental Watershed (DCEW) we applied STAR-FM in an ecosystem that exhibits plot (e.g., $10^0$-$10^1$ m), hillslope (e.g, $10^1$-$10^2$ m) and watershed (e.g., $10^2$-$10^3$ km) scale heterogeneities. DCEW is associated with grasses at lower elevations, mixed shrubland ecosystems at mid-elevations, and coniferous evergreens at higher elevations. At plot scales, heterogeneity arise in shrubland areas within the watershed with sagebrush interspersed with bare ground and perennial grasses. At hillslope scales, heterogeneity arises because of contrasts in vegetation arising due to contrasts in incoming solar radiation that is associated with hillslope aspects. At lower elevations, shrubs can be present on North-facing aspects while grasses predominate on South-facing aspects. At middle elevations, trees begin to become common on North-facing aspects, whereas shrubs are more prevalent on South-facing aspects. At watershed scales, the transition from grasses at lower elevations to coniferous evergreens at higher elevations is associated with a corresponding gradient in precipitation.

We developed and conducted a series of experiments to test the accuracy of the STAR-FM framework in interpolating high-temporal resolution MODIS data in space to spatial resolutions of Landsat. The STARFM fusion approach uses one or more pairs of Landsat (acquired every 14 days[NG1] ) and MODIS (acquired daily) surface reflectance images to predict daily surface reflectance at a Landsat spatial resolution (30 m) for dates where no Landsat images were acquired. The model assumes aggregation from finer scale homogenous pixels into coarser heterogeneous pixels and is therefore sensitive to landscape patch size and degrades when applied to landscapes with fine scale heterogeneity. The STARFM model produces a synthetic Landsat image from a temporally coincident pair of Landsat and MODIS images, and an additional MODIS image on the prediction date.

We assessed model accuracy with a pixel-based differencing of the STARFM predicted and the observed Landsat images of NDVI and individual green, red, and near-infrared bands . We calculated mean absolute error (MAE), mean signed error (MSE), standard deviation of the error, and root mean square error (RMSE). Linear regression analysis was performed for all pixels in the image to determine bias (slope and intercept) and overall prediction accuracy ($R^2$). Finally, we provided the Nash-Sutcliffe Efficiency index (NSE) as an alternative to traditional regression-based analyses by comparing the prediction error with a mean value (McCuen et al. 2006). The NSE index varies from $-\infty$ to 1, where a value of 0 indicates no improvement over the mean, negative values signify poorer prediction than the mean, and 1 is a perfect match to observed data.

Inputting two pairs of training data, one pair before the predicted date and one pair after, into STARFM resulted in a more accurate synthetic NDVI image than using only one pair of training data (**Figure 1**). Early spring had the highest model errors, particularly a late-April image (4/27). Key finding #3: Despite poorer results early in the growing season, all synthetic images in the two-pair model contained new information (NSE > 0) and images beginning in late May through the end of the growing season explained over 94% of the variation.

In both cases where two images (one before and one after the date of interest) and one image (one before the date of interest) were used to train the STARFM algorithm, as evidenced by positive values of NSE, the algorithm was able to add value to the understanding of the spatiotemporal dynamics of vegetation in these semiarid ecosystems. This added information is potentially useful in the context of validating or constraining spatially distributed models at hillslope scales. In particular, it provides an indirect/weak constraint on the behavior of soil moisture in the root zone during the period of interest. It is possible that application of the STARFM fusion model to semiarid, heterogeneous watersheds could be improved, especially in the “shoulder months” of the growing season, by using a MODIS product with higher temporal resolution than the 16-day NBAR composite Nadir.

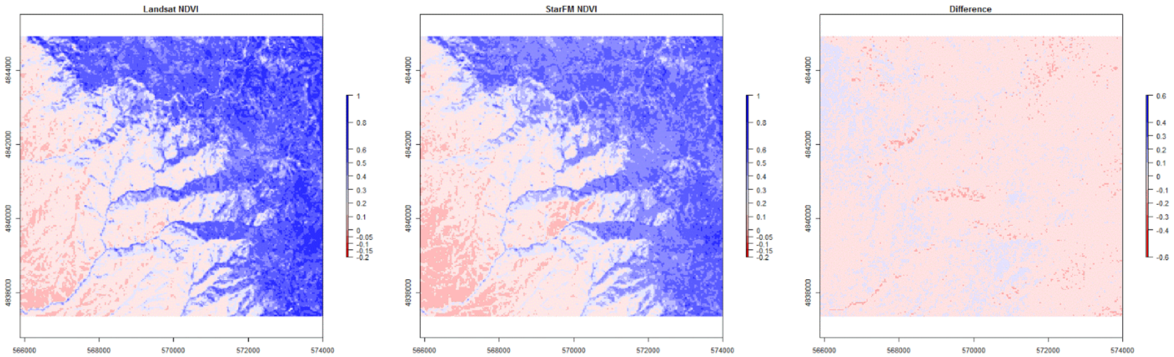

**Figure 1:** The Landsat observed (left), STAR-FM simulated (center) and difference (right) in normalized difference vegetation index (NDVI) over the Dry Creek Experimental Watershed on 30 June 2007.

We are developing a manuscript summarizing these experiments and the associated implications for ecohydrologic modeling are discussed. This manuscript will be submitted for publication to a peer-reviewed journal within the next month.

### *2.2 Fusing multitemporal remote sensing imagery with ground-based observations*

A complimentary line of inquiry that was pursued during the grant investigated fusing multi-temporal remote sensing imagery with ground-based observations of net ecosystem exchange in a heavily instrumented experimental watershed. The overarching objective of this component of the work is to create a multiple-data source estimate of net ecosystem exchange in space and time, as well as a measure of the uncertainty in that estimate. We hypothesize that such a data product could serve as a valuable constraint on estimates of evapotranspiration and net primary productivity predicted from physics-based ecohydrology models.

### *2.2.1. Study area*

This component of the research made opportune use of a suite of eddy flux towers measuring land-atmosphere exchanges of water, carbon, and heat in the Reynolds Creek Experimental Watershed (RCEW) and Critical Zone Observatory (RCEW). RCEW is an approximately 239 km$^2$ watershed in the Owyhee Mountains in southwest Idaho, USA. The watershed was established in the 1960s by the US Department of Agriculture Agricultural Research Service (ARS). Elevation in RCEW ranges in elevation from 1100 to 2240 m. The southern portion of the watershed is associated with the uplands and Reynolds Creek drains to the north toward the Snake River Plain. Vegetation in at lower elevations in RCEW is consistent with the sagebrush-steppe ecosystem, sagebrush interspersed with perennial grasses and bare soils with significant biotic crusts. In recent years, invasion of annual grasses (e.g., B. tectorum) has become more common in the lower elevations of the watershed. Predominant vegetation patterns shift in the higher elevations of RCEW. Trees common in higher elevations include both coniferous evergreen trees, as well as isolated Aspen stands that are typically found in regions of convergent topography. Mountain sage communities, adapted to cooler temperatures, are also found at higher elevations in RCEW, with ceanothus shrubs common

Mean annual precipitation in RCEW ranges from 200 mm/yr at the bottom of the watershed to more than 1100 mm at the top [Hanson, 2001]. The semiarid climate is associated with a cool, wet winter and hot, dry summer. During the wet season rainfall dominates at lower elevations in the watershed while snowfall dominates in the upper elevations. Previous studies have documented the importance of snow drifts on water cycle dynamics in RCEW [Marks and Winstral, 2001]. Throughout its history RCEW has served as an important facility for the study of rangeland hydrology and ecology, fire science, and geomorphology.

Since 2007 5 eddy covariance towers have operated on a mostly continuous basis. These eddy covariance towers were initially situated in and around Reynolds Mountain East in the southern, high-elevation portion of the watershed and were installed to investigate sublimation from winter snowpacks. Since the advent of the CZO in RCEW, three of these have been redeployed in order to better characterize ecosystem-atmosphere exchanges of water and carbon. In this work we examined two years of eddy covariance Net Ecosystem Exchange (NEE) data at the Aspen and High Sage sites.

### *2.2.2 Methods*

Our objective for this portion of the research was to develop a statistical model to predict NEE using data that could, in principle, be derived from remote sensing information only. Potential predictor variables include: (1) topography and associated derived variables like elevation, aspect, slope, and potential incoming solar radiation, (2) terrestrial vegetation information in the form of vegetation indices from the MODIS sensor, (3) soil moisture, and (4) antecedent precipitation.

A random forest approach was used to derive an ensemble of classification and regression trees. In the random forest approach, the model building process works as follows:

1. The predictor and response variable are segmented into data for training (so-called in-bag samples) and data for verification (so-called out-of-bag samples)
2. A regression trees is built by fitting the tree to the in-bag sample
3. Procedures such as tree pruning can be used at this stage of model development
4. Error metrics (e.g., root mean squared error, etc.) are computed by using the developed regression tree to predict the values of the out-of-bag sample.
5. The regression tree and error metrics are saved and the process of segmenting the dataset into in-bag and out-of-bag samples, model development and verification repeat

**Table 1. Variable importance based on the RF model**

| Predictors | Importance of predictors |
|---|---|
| fPAR | 1.091 |
| LAI | 1.673 |
| Precipitation | 0.205 |
| Mean Air Temperature | 1.599 |
| Downward Solar Radiation | 1.674 |
| Soil Moisture | 1.665 |

With this ensemble of regression trees, the importance of predictor variables can be inferred by developing a histogram that characterizes how frequently the predictors appear in the the developed regression trees (**Table 1**). Additionally, the predictability of the response variable can be assessed by investigating the distribution of the error metrics for the ensemble of regression trees. The random forest approach – a machine learning (i.e., "big data") methodology – is a straightforward and useful way to develop predictive models of important ecohydrologic variables like net ecosystem exchange (**Figure 2**). We developed a predictive model, moreover, that is capable of using as input variables that are available via remote sensing. Of particular importance is the role of soil moisture, air temperature, and solar radiation as predictors of NEE. This initial effort was focused on evaluating and validating models at a particular site. Future efforts will investigate the extent to which these models are transferrable to similar locations where ground-based eddy flux data is withheld in a data-denial experiment. This will provide insight into whether these machine learning approaches are viable for predicting key battlespace environmental variables in locations of limited and denied access.

### *2.3 Developing spatially continuous, serially complete, internally consistent environmental forcings*

One of the biggest remaining challenges with predicting soil moisture and (more broadly) hydrologic forecasting at spatial scales of importance to the Army in topographically complex

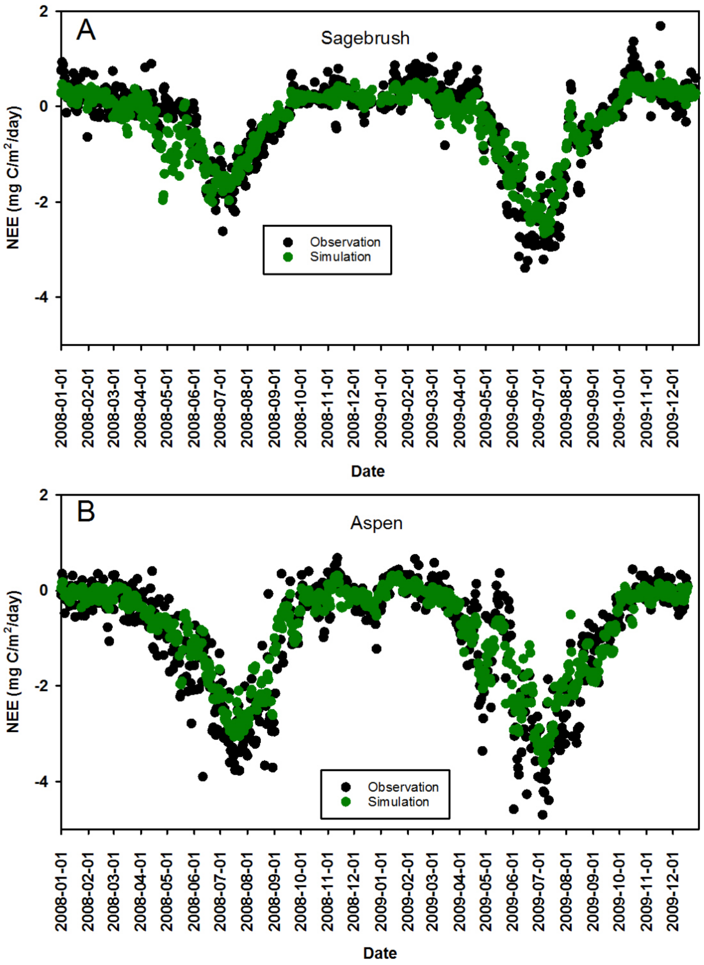

**Figure 2:** Daily values of net ecosystem exchange from the random forest model (green circles) and observations (black circles) for the (a) sagebrush and (b) aspen sites.

terrain remains the lack of hydrometeorological forcings at adequate spatiotemporal resolutions. Such data are needed to develop, parameterize, calibrate, verify, and assess uncertainty in hydrologic models capable of resolving hillslope-scale variation in soil moisture and other important hydrologic variables. Available globally available, operational data products are far too coarse to capture important meteorological phenomena in mountain watersheds, particularly orographic forcing that leads to large precipitation lapse rates.

The National Center for Environmental Prediction's (NCEP) Global Forecasting System (GFS) provides global meteorologic forecasts in the 0- to 10-day time horizon every six hours, but at spatial resolutions of 0.5 degrees (approximately 56 km at the equator). NCEP also provides their GFS Final Analysis (GFS-FNL) data product that is a near-real time estimate of land and atmospheric conditions, constrained to observational data, also at spatial resolutions of 0.5 degrees. Using the same modeling framework as the GFS model, NCEP's Climate Forecasting System version 2 (CFSv2) provides extended-range climate forecasts in the 0- to 9-month time horizon at 0.5 degrees. NCEP also provides data from the Global Data Assimilation

System (GDAS), a reanalysis product that blends both models and a suite of observational data, to support retrospective analysis. GDAS data are associated with a spatial resolution of as fine as 0.25 degrees (28 km at the equator) and hourly temporal resolutions.

While invaluable for hydrologic modeling at continental scales, the above datasets do not provide adequate spatial resolution for modeling the hydrology of large watersheds at spatial scales of interest for Army applications. A complementary track of the research, which was made possible by joint support from additional grants awarded during this ARO project, investigates the use of a regional climate and weather forecasting tool for generating environmental forcings at spatial resolutions sufficiently fine to use as input to hillslope-scale resolving hydrology models.

The regional climate and weather forecasting tool used, the Weather Research and Forecasting (WRF) model [Skamarock et al., 2008], was developed as a partnership between many federal agencies and research centers, particularly the National Center for Atmospheric Research (NCAR). The WRF model simulates the coupled dynamics of the land-atmosphere system over a region of interest, taking forecast or reanalysis data (e.g., GFS or GFS-FNL data) as input. These input data are used to prescribe atmospheric conditions like temperature, humidity, pressure, and wind vector at the lateral boundaries of the simulation domain and land surface conditions like soil moisture and temperature at the initial simulation time within the domain.

During a forecast or hindcast simulation, the WRF model simulates both atmospheric and land surface hydrology variables in response to these inputs. Effectively, this approach embeds a user-defined, finer-resolution simulation domain within these coarser forecast and reanalysis products. Key atmospheric variables include surface precipitation, temperature, humidity, pressure, wind speed and direction, and incoming solar and longwave radiant fluxes. Land surface hydrology variables simulated include soil temperature and moisture, snow water equivalent, and runoff. Importantly, provided that these atmospheric variables are simulated at sufficiently fine spatiotemporal resolutions using the WRF model, they can be used as input to hydrologic models capable of simulating hydrologic response at spatial scales of Army interest. A key feature of the WRF model allows for nesting of higher resolution domains within the primary domain. WRF-simulated atmospheric conditions are used to prescribe the boundary conditions for the higher resolution domain. This feature allows the user to specify multiple domains within a region to "step down" the spatial resolution of simulated land-atmosphere conditions to a resolution appropriate to the application and area of interest. This "nesting" feature can be used in a one-way or two-way coupling model. In one-way coupling the boundary conditions from WRF are used to prescribe boundary conditions in the higher resolution domains within the primary domain. In two-way coupling the land-atmosphere conditions simulated in the interior domains are aggregated to the resolution of the parent domain and used in place of the simulated conditions in that coarser domain. This two-way coupling allows for the influence of higher resolution heterogeneities in land surface variables like topography and vegetation to propagate out to the coarser domain.

In order to develop a benchmark dataset for future hydrological research and serve as a baseline assessment of the ability of models like WRF to simulate hydrometeorological conditions at climate timescales, we are creating a 30-year historical climate reanalysis over the interior Pacific Northwest region. The outer domain is simulated at a spatial resolution of 3 km, with outputs from the WRF model saved at 3 hr intervals while the inner domain is simulated at a spatial resolution of 1 km, with outputs saved at 1 hr intervals (**Figure 3**). The outer domain contains the entire Snake River basin, while the inner domain contains the entire Boise River basin and some key areas of interest extending outside the Boise River watershed. Key areas included in the inner domain include RCEW, Mountain Home Air Force Base, and the Orchard Combat Training Center (OCTC).

As input, the simulation uses historical reanalysis data in the form of the North American Regional Reanalysis (NARR). The NARR dataset is available from 1979 onward at spatial resolutions of 0.3 degrees (32 km at the equator) and temporal resolutions of 3 hr. NARR was developed by NCEP and is a data assimilation product that combines model output with important atmospheric observations, notably including precipitation. Simulations are conducted on a water year basis, beginning on October 1 of each year. The simulation begins on October 1, 1984 (the 1985 water year) and continues through water year 2015.

We performed the simulations on the Yellowstone supercomputer at the NCAR Wyoming Supercomputing Center using a large allocation granted through EPSCoR. Uncompressed the dataset is 140 TB in size and Boise State Research Computing has provided adequate storage space, leveraging the relatively new features of data de-duplication, to realize a 40% compression. This means that the dataset can be stored on 80 TB of actual storage space, but accessed without significant delay owing to the nature of the data de-duplication process. Surface

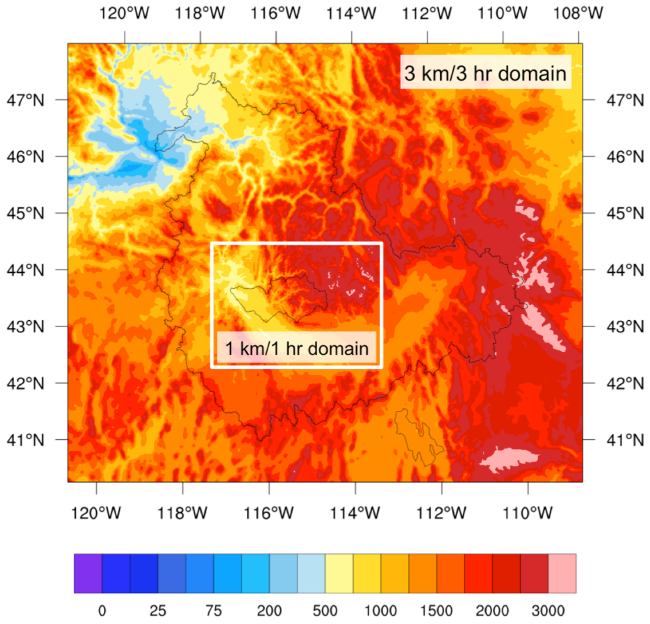

**Figure 3:** WRF domains for the 30 year, long term simulation. The outer 3 km domain encompasses the entirety of the Snake River Basin.

hydrometeorological conditions are freely available via an OPeNDAP web portal set up by the Boise State Research Computing. The OPeNDAP server allows users to retrieve subsets of the entire dataset in space, by variable, or both. The Boise State library created a digital object identifier (DOI) for the dataset and it can now be cited in scholarly works.

We are currently analyzing the dataset against a network of surface observations, remote sensing products, and complementary geostatistical interpolations of surface observations. Once this analysis is complete we will submit a manuscript summarizing this dataset, the error characteristics, and summarizing potential future applications. To our knowledge, this is the longest climate-scale simulation available for such a large watershed and at such a high spatial resolution.

Research activities we will undertake in the near future will include efforts to reduce errors in the dataset, removing any bias and constraining surface hydrometeorological variables – particularly those variables that would be used as inputs to hydrological models – to observations. This dataset and its derivatives will serve as an important legacy of this ARO Young Investigator Program grant. It will serve as an important forcing dataset to models of land surface/subsurface hydrology and terrestrial ecology. Because of the spatiotemporal completeness, and high resolution of this dataset the following research avenues are areas that our research group has already begun to explore or will explore in the near future:

***Snow remote sensing and data assimilation:*** Although a number of proposals have been put forward for a snow remote sensing platform, there is still no general consensus about the sensor package, spatial resolution, and temporal revisit characteristics of such a mission. Lidar and multi/hyperspectral remote sensing platforms can provide information about snow depth and extent but cannot provide direct measurements of snow density or water content. At the same time, microwave remote sensing is potentially useful for retrieving information about snow density and water equivalent, but snow retrieval algorithms and radiative transfer modeling remains in its infancy. Platforms that combine lidar, multi/hyperspectral, and microwave sensors could potentially leverage the strengths of each technology while ameliorating the associated shortcomings of each. The cost and mass of such a payload, however, might render such a mission prohibitively expensive. The 30 year WRF dataset provides an important opportunity to engage in the snow remote sensing community. This dataset, when used in conjunction with hillslope-scale resolving models of snow accumulation and ablation, will enable a suite of synthetic data assimilation experiments to address important outstanding research questions in snow remote sensing. Such questions include: For a given desired basin-scale snow water equivalent retrieval error, what is the required accuracy of snow water equivalent from satellites? What is the required spatial resolution and temporal revisit of a satellite given these accuracy constraints? How do errors in retrieved snow water equivalent propagate through the hydrologic system to produce errors in runoff and streamflow?

***Remote sensing of terrestrial ecosystems:*** The current generation of multispectral satellites (e.g., MODIS, Landsat) has yielded significant capabilities to monitor the long-term dynamics of terrestrial ecosystems. These technologies, however, provide only limited information about both the structure and function of these ecosystems. New and emerging technologies, however, show significant promise in advancing the characterization of both ecosystem structure and function from space. Lidar remote sensing, specifically, can provide data about ecosystem structure, potentially including canopy height, volume, density, and intercanopy spacing. Hyperspectral data provides data on the spectral response of terrestrial vegetation that much more continuous in the frequency domain than existing multispectral satellites. Hyperspectral remote sensing techniques have show promise in providing details about terrestrial vegetation like foliar nitrogen content, carbon-nitrogen ratios, and stress [Mitchell et al., 2012]. Lidar and hyperspectral remote sensing techniques, therefore, can provide potentially information about terrestrial ecosystems that is of significant value to models of the terrestrial ecosystem. The 30-year environmental forcing dataset will support new efforts to advance the accuracy of ecohydrological forecasting and data assimilation in semiarid systems. Specifically, using airborne lidar and hyperspectral remote sensing assets we can optimize characteristics of potential satellite missions to characterize terrestrial ecosystems like spatial resolution, spatial revisit, and sensor properties. We can also investigate how accurately characteristics of terrestrial ecosystems that cannot be directly observed, such as soil organic matter and litter-fall, can be retrieved through a data assimilation framework.

***Assessing impacts of land management in a changing climate:*** The 30 year WRF dataset is also supporting new directions of research aimed at investigating the extent to which land management activities like fuels treatment in forests and rangelands, selective thinning, and ecosystem restoration impact climate at regional scales. As a historical reanalysis of climate produced at high spatiotemporal resolution, the 30-year dataset serves as a benchmark against which simulations of alternative scenarios of land management can be compared. Since the beginning of the grant, the PI has developed agent-based modeling (ABM) skills as a way of parameterizing autonomous behavior of land management agencies in public lands throughout the western US. ABMs can be used to simulate the effects of management actions on terrestrial vegetation under a variety of policy actions. These altered vegetation patterns can then be fed back into the WRF model to assess the associated impacts of vegetation change on regional climate and hydrology. Such scenario analyses can be conducted under both historical and projected climate conditions to quantify how land management actions may have influenced regional climate and hydrology in the past, as well as the impact that those actions might have in a future climate. These modeling studies can thus be used to inform land management agencies and personnel about impacts of management actions on regional hydroclimate now and in the future.

***Coupled modeling of food, energy, and water systems (FEWS):*** The development of this dataset occurs at a particularly germane time, given new interest in the sustainability of the nexus of food, energy, and water systems (FEWS) by the National Science Foundation, the Department of Agriculture, US Geological Survey, and the Department of Energy. As a dataset that represents a "best available estimate" of weather and climate in the region during the previous 30 years, this dataset also serves as an important resource for integrated analysis of the food, energy, and water system in the semiarid interior Pacific Northwest. The Snake River Basin and the Boise River Basin represent large watersheds where water is a vital resource for both production of food and hydropower. We anticipate that the availability of this dataset will ignite new research pathways in modeling FEWS in an integrated way that encompasses key elements of both biophysical and social systems. During my upcoming sabbatical during the 2016-17 academic year, I will in particular be focusing on developing a portfolio of FEWS-related research questions. Some key questions that will be addressed during this period include: What is the appropriate scale at which to address the sustainability and resilience of FEW systems? Along what social and biophysical dimensions should the resilience and sustainability of FEW systems be assessed? And, what are the key biophysical and social datasets and variables required to support such an assessment of FEW system sustainability and resilience?

***Accuracy of extended-range weather forecasts in complex terrain:*** A key enterprise for both water resource and Army applications is forecasting weather and hydrometeorological conditions in the battlespace environment in the extended time-frame (e.g., 10-day to 6 months). This time horizon is of particular importance because of the logistical challenges associated. Often, months are required to position troops and materiel in advance of Army operations. It is therefore desirable to have some information, as well as ancillary understanding of the accuracy of that information, about potential battlespace environment conditions during the window of time being considered for these operations. The 30-year dataset provides an important benchmark against which experimental extended range, high-resolution forecasts can be assessed. As previously mentioned, because it integrates data assimilation datasets that contain information from a diverse suite of observational assets, this 30-year WRF dataset represents the “best estimate” of historical hydrometeorological conditions during this time window. NCEP has archived reforecast data for the CFSv2. These reforecast data provide historical forecasts from a date of interest out to 6 months after that initial date. Using these reforecast data as input to the WRF data, we can assess the degree to which regionally focused, dynamically downscaled CFSv2 forecasts could have reproduced historical weather, as captured by the WRF 30-year reanalysis. This research will lead to both enhanced fundamental understanding of land-atmosphere interactions in complex, mountain watersheds. It will also yield practical understanding about how to configure resources like the WRF model in the context of CFSv2 data to produce extended range forecasts that are useful in the context of both civil and Army decision-making.

## 3. Products

### *3.1 Completed*

Kormos, P. R., J. P. McNamara, M. S. Seyfried, H.-P. Marshall, D. Marks, and A. N. Flores (2015), Bedrock infiltration estimates from a catchment water storage-based modeling approach in the rain snow transition zone, Journal of Hydrology, 525, 231-248, doi:10.1016/j.jhydrol.2015.03.032.

Walters, R. D., K. A. Watson, H.-P. Marshall, J. P. McNamara, and A. N. Flores (2015), A physiographic approach to downscaling fractional snow cover data in mountainous regions, Remote Sensing of Environment, 152, 413-425, doi:10.1016/j.rse.2014.07.001.

Kormos, P. R., D. Marks, J. P. McNamara, H.-P. Marshall, A. Winstral, and A. N. Flores (2014), Snow distribution, melt and surface water inputs to the soil in the mountain rain-snow transition zone, J. of Hydrol., 519(A), 190-204, doi:10.1016/j.jhydrol.2014.06.051.

Flores, A. N., D. Entekhabi, and R. L. Bras (2014), Application of a hillslope-scale soil moisture data assimilation system to military trafficability assessment, J. Terramechanics, 51, 53-66, doi: 10.1016/j.jterra.2013.11.004.

### *3.2 Products Still in Progress*

Flores, A. N; M. Masarik, and K. A. Watson (2016), A 30-Year, Multi-Domain High-Resolution Climate Simulation Dataset for the Interior Pacific Northwest and Southern Idaho, The Lab for Ecohydrology and Alternative Futuring (LEAF), Dataset 1, doi:10.18122/B2LEAFD001.

Zhou, Q., A. N. Flores, A. Fellows, and G. Flershinger, Deriving temporal net ecosystem exchange estimates by combining eddy flux, remote sensing and WRF model data, manuscript in prep.

Olsoy, P.J., A.N. Flores, J. Mitchell, and N.F. Glenn, Spatiotemporal monitoring of vegetation patterns in heterogeneous, dryland ecosystems using remote sensing: assessing the efficacy of a multi-platform data fusion technique, manuscript in prep.

Zhou, Q., A.N. Flores, R.D. Walters, and B. Han, A machine-learning approach to estimation of downward solar radiation from satellite-derived data products: an application over topographically complex terrain in U.S., manuscript in prep.

### *3.3 Additional Grants Enabled by Work*

This YIP award has contributed to substantial early career success in grantsmanship by the PI. The work undertaken during this award has catalyzed new collaborations, new partnerships with agencies, professional development opportunities, and new research directions in coupled land-atmosphere modeling, remote sensing, and data assimilation. These grants address the key outgrowths of research identified in the previous section, as well as new opportunities that arose during the course of this grant. The awards most directly related to work enabled by this ARO YIP grant are itemized below:

1. CAREER: Citizens, Conservation, and Climate: Research and Education for Climate Literacy in Managed Landscapes; Role: PI; Source: National Science Foundation; Performance period: June 1, 2014 – May 31, 2019; Total budget: $457,205.
2. Multiple frequency active microwave remote sensing for snow water equivalent retrieval from space: a data assimilation approach; Role: PI; Source: NASA Terrestrial Hydrology Program; Performance period: July 1, 2015 – June 30, 2018; Total budget: $295,577.

3. Monitoring Earth's Hydrosphere from Space; Role: Science PI; Source: NASA EPSCoR; Performance period: November 1, 2014 – October 31, 2017; Total budget: $750,000.
4. Intermediate-range Climate Forecasting to Support Water Supply and Flood Control with a Mesoscale Model; Role: PI; Source: US Bureau of Reclamation; Performance period: May 1, 2015 – April 30, 2018; Total budget: $150,000.
5. Scalable vegetation structure for ecosystem modeling in the western US; Role: Co-I; Source: NASA Terrestrial Ecology Program; Performance Period: January 1, 2014-December 31, 2016; Total budget: $748,916.
6. Reynolds Creek Carbon Critical Zone Observatory; Role: Co-PI; Source: National Science Foundation; Performance period: January 1, 2014-December 31, 2018; Total budget: $2,500,000.
7. EPSCoR RII Track 2: Collaborative Research: The Western Consortium for Watershed Analysis, Visualization, and Exploration (WC-WAVE); Role: Co-PI; Source: National Science Foundation; Performance period: September 1, 2013-August 31, 2016; Total budget: $2,000,000.

## 4. Future Directions and Implications for Army Science

The research accomplished during this grant will drive future science efforts that are remain particularly relevant to Army applications. A few of the near-term applications of interest are provided below.

### *4.1 Improved Landslide Forecasting*

In mountainous regions of the world (e.g., the Hindu Kush Mountains), movement of Army supplies, troops, and materiel is often confined to valley-bottoms that may or may not contain roads of varying quality. If sufficiently steep, the hillsides abutting these roads can be prone to rockfalls, landslides, debris flows, and – when snow is present – avalanches. Instability of these slopes can also be exacerbated by poor controls on land use in the region, disturbance to vegetative cover (e.g., fire), and antecedent soil moisture and precipitation. These mass movements – if they occur during Army operations – can be a significant threat to the success of these operations and lives of operators.

At present, there is a lack of predictive frameworks that leverage the availability of multi-scale datasets to predict the susceptibility of landslides and other mass movements in a given region at spatial scales of interest to the Army (10s to 100s of m). The research activities performed during and catalyzed by this YIP grant set the stage to address this important knowledge and information gap.

Specifically, this YIP grant has led to advances in the modeling of coupled land-atmosphere models at high spatial resolutions in complete terrain, improved abilities to extract information from remote sensing datasets, and the use of machine learning to improve prediction and forecasting in the hydrologic sciences. The problem of quantitatively assessing the risks of landslides at a particular location of interest can be reduced to a workflow that involves: (1) developing hydrometeorological forecasts at appropriate spatiotemporal resolutions in the area of interest using a model like WRF, (2) using those forcings, with knowledge of topography, vegetation, cover and soil characteristics, to derive spatiotemporal distributions of soil moisture and pore pressure at appropriate spatial resolutions and with adequate lead time, (3) a means of quantitatively assessing local slope stability given the local topographic and soil properties, and simulated soil water/pore pressure status, and (4) an effective methods to visualize and communicate this risk and the uncertainty in that assessment.

The research highlighted above addresses key aspects of this workflow. In particular, the research accomplished in this grant makes strides toward improving the ability to parameterize and use hydrologic models for applications like landslide susceptibility analysis in generic and complex landscapes.

### *4.2 The role of agricultural systems in Army operations*

Agricultural land uses, particularly irrigated areas, impact Army operations in important and multi-faceted ways. Irrigation tends works to increase the latent heat flux from the land surface, having a net cooling impact at the surface. The presence of irrigated areas in desert landscapes, therefore, tends to result in a large contrast in both the partitioning of turbulent heat fluxes into latent and sensible heat, as well as in the radiant surface temperature. These large contrasts can create significant problems with night vision equipment, particularly when the skin temperature of the land surface is close to that of the ambient atmosphere. This effect, known as thermal crossover, can lead to an inability to distinguish between the land surface and atmosphere resulting in hazards to operators and operations. Contrasts in turbulent fluxes, moreover, can create problems for low altitude aviation operations – particularly for rotary winged aircraft and unmanned aerial systems. These can lead to corresponding contrasts in circulations within the boundary layer and surface wind shear that can pose safety-of-flight challenges.

For these reasons, it is critical to improve methods for predicting the hydrometeorological response at the land surface and in the planetary boundary layer in regions with large contrasts in land use that result in corresponding heterogeneity in the partitioning of latent and sensible heat. This research project, and ongoing research projects being conducted in the PI’s research group, could contribute to enhanced predictions of near-surface hydrometeorology in these regions in a number of ways.

First, the spatial resolutions needed for resolving sharp transitions from desert ecosystems to irrigated agriculture are on the order of 10s of m, at most. To simulate the coupled dynamics of the land surface and planetary boundary layer a large eddy simulation (LES) parameterization of the boundary layer needs to be used. Although this increase in resolution leads to significant increases in computational cost, the improved spatial representation provides new opportunities to ingest novel sensor data. Specifically, thermal infrared data collected from unmanned aerial systems (UASs).

Our group is beginning preliminary simulations with the LES parameterization in WRF to simulate coupled land-atmosphere systems at spatial resolutions approaching 30 m. From these initial experiments we will be developing some preliminary datasets and studies to suggest how data and models might be used together in these heterogeneous environments to improve predictions of the distribution of latent and sensible heat fluxes, applied to improved characterization of environmental conditions during Army operations.

### *4.3 Extended range trafficability assessment for Army applications*

As previously stated, extended-range forecasting is important for both water resources and Army operations. Knowledge of hydrometeorologic conditions in the battlespace environment in advance of an operation is extremely helpful for planning and conduct of military missions. Often, months are required to position troops and materiel in advance of Army operations. It is therefore desirable to have some information, as well as ancillary understanding of the accuracy of that information, about potential battlespace environment conditions during the window of time being considered for these operations.

In support of a grant from the Bureau of Reclamation Science and Technology program, our lab is investigating the ability of the WRF model, informed by the previously mentioned CFSv2 data, to predict hydrometeorological forcings in the 0-30 day ahead time. For the 3 km WRF Snake River domain described above, we will be developing weekly forecasts with the WRF-CFSv2 model for Water Year 2017 (1 October 2016 – 30 September 2017). We will dynamically downscale all four of the CFSv2 ensemble replicates produced on the Thursday 00z forecast. This exercise will allow us to quantify not only the potential environmental conditions up to 30 days in advance, but also a measure of the spread in the forecasts themselves.

The Orchard Combat Training Center (OCTC) is located within this domain. OCTC is an important desert training environment for Army National Guard units around the country. It is equipped with state-of-the-art gunnery for a variety of weapons training, including for small arms, crew served weapons, Bradley, Abrams, field artillery, mortars, and helicopter gunnery. The PI has begun reaching out to staff at OCTC in order to potentially collaborate on a future submission to investigate the efficacy of extended range forecast products for assessing range mobility and trafficability.

## REFERENCES

Gao F, J. Masek, M. Schwaller, and F. Hall (2006), On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance, IEEE Transactions on Geoscience and Remote Sensing 44(8):2207-2218. doi:10.1109/TGRS.2006.872081.

Hanson, C.L. (2001), Long- term precipitation database, Reynolds Creek Experimental Watershed, Idaho, United States. Water Resources Research, 37(11), 2831-2834.

Marks, D., and A. Winstral (2001), Comparison of snow deposition, the snow cover energy balance, and snowmelt at two sites in a semiarid mountain basin, Journal of Hydrometeorology, 2(3), 213-227.

McCuen, R.H., Z. Knight, and A.G. Cutter (2006) Evaluation of the Nash–Sutcliffe efficiency index, Journal of Hydrologic Engineering, 11(6), 597-602.

Mitchell, J. J., N.F. Glenn, T.T. Sankey, D.R. Derryberry and M.J. Germino (2012), Remote sensing of sagebrush canopy nitrogen, Remote Sensing of Environment, 124, doi:10.1016/j.rse.2012.05.002, 217-223, Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. Duda, X.-Y. Huang, W. Wang and J. G. Powers (2008), A Description of the Advanced Research WRF Version 3, NCAR Tech. Note TN-475+STR, 113 pp.