REVIEWER # 5:

1) Parts of the background, i.e. the effects of each of the input in the TWI formulation is presented and accounted for in the methods section. This is unevenly distributed with rudimentary background/discussion on cell size and slope calculations as oppose to the flow accumulation algorithm. The benefits and downsides of the other inputs would strengthen the discussion in the end.

   RESPONSE: We have added additional description of cell size and slope calculation methods to the Methods section. Please refer to lines 252-259 and 268-277.

2) The resulting derivatives of elevation models are to a large extent a product of the input, in this case the datasets called USGS and LiDAR. The data quality of those datasets differs with RMSEs of 2.44m and 0.15m respectively. An evaluation of these parameters or at least an overview of what is stated in the literature on this topic would improve the paper.

   RESPONSE: We now discuss this topic in lines 527-533. We have added more explicit acknowledgement that the USGS and LiDAR DEMs are using data of inherently different quality (section 3.1). Indeed, this issue is exactly why we do not directly compare TWI analyses between USGS and LiDAR throughout much of the paper; we now explain this explicitly (lines 462-465). We think the issues of data quality between these LiDAR-based and photometric-based DEMs has been the specific focus of research elsewhere. We have cited some representative studies that are most relevant to this study.

3) When it comes to the smoothing scheme carried out in the study, the problem of TWI's with "high local variations" is something that concerns me. It is not explained how this method logically would improve the result. Smearing out errors in a DEM with a RMSE of 2.44m generate a different impact on elevation values than in a DEM with a RMSE of 0,15m. This issue is more of a data quality issue than how to obtain a good TWI value. I argue that if such a procedure as filtering is needed then the data is not at all optimal for the task. Another issue related to this is if DEMs were filtered before all derivatives were calculated?

   RESPONSE: Only the final TWI maps were filtered, not the raw DEMs. The RMSE in the LiDAR data refers specifically to the point cloud data - not the interpolated DEM (this is now specifically noted in lines 241-242). Our findings did indicate that in the case of LiDAR TWIs, filtering was unnecessary so data-quality vs. filtering was a non-issue here. We have noted that the smoothing is off-setting the consequences of using lower quality USGS DEMs (line 160). Regarding the logic of smoothing TWIs we present a thorough discussion in section 3.6 (lines 629-637). However, it is also worth emphasizing that the "high local variations" refer specifically to high variation in TWI values between a cell and its neighboring cells. This can lead to situations where high TWI values (wet areas) in flow convergent areas are unrealistically concentrated/linear with exceedingly large contrasts between neighboring cells. Please refer to Figure C (below) for an illustration of this phenomenon; note in

particular the dark blue cells (wet) immediately adjacent to very dry cells (dark red). In reality, near-stream areas can exhibit high soil moisture contents due to a combination of return flow and groundwater mounding proximal to effluent streams. So in many instances, the transition from channel cells in valley bottoms to drier upslope cells should be more gradual than represented by some TWIs. Filtering a TWI may help to smooth out this transition – thereby potentially rendering it more realistic.

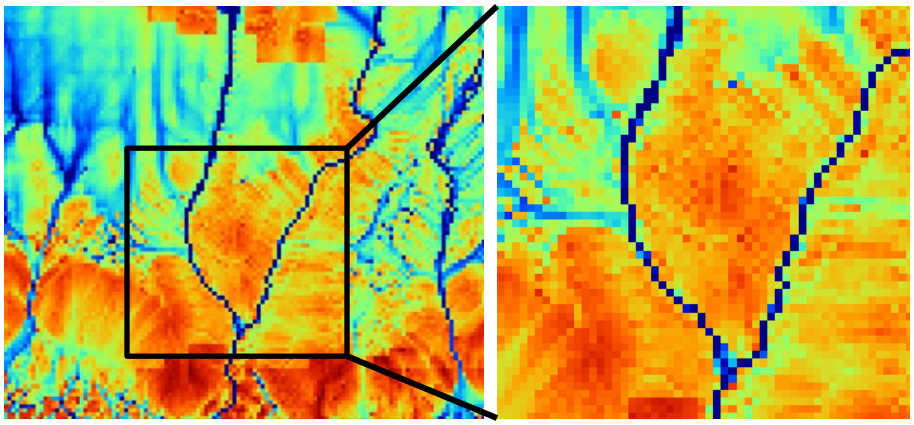

Figure C. Example TWIs demonstrating "high local variation" in TWI values. Dark Blue and dark red pixels represent wet dry areas, respectively.

4) The source data discussion is also related to the data quality and resolution. Since LiDAR data often offer a higher point density and a better point distribution, the generation of DEM with almost any chosen interpolation algorithm is better than data obtained with a coarser sampling method. A deeper discussion of the data quality, generation of DEMs and resolution on how these parameters affect the TWI in the study should be added. The connection between the quality of data and the generation of TWI is also missing in the flow accumulation discussion. I get the impression that the authors want to connect this with the sampling method. The way the data was obtained is only important in the data quality, point density, point distribution and DEM generation. This applies to both the LiDAR and the USGS discussion.

RESPONSE: We discussed this issue in lines 438-457 in section 3.1. In addition, we have added additional text which further explores this topic. Please refer to lines (518-520 and 526-532).

5) In the section 3.5.1, P14060L12- 16, the authors obviously think there is a risk that conversion may occur in their dataset as Park's results are referred to. However, it is also stated that this was for larger cell sizes than 20m. In this study the resolutions studied are 3 and 10m so why is there a need for this in the present study?

RESPONSE: Both Park et al. (2009) and Erskine et al. (2006) found that relative differences between SFD and MFD algorithms were inversely related to cell size. Park et al. (2009) evaluated 1, 2.5, 5, 7.5, 10, 15, 20, 30, 40, 50 m grid cell sizes, while Erskine et al. (2006) evaluated 5, 10 and 30m. So, their range of tested cells sizes encompasses ours. However, they did not evaluate LiDAR-derived DEMs so we felt it was prudent to investigate whether

there was an interaction between flow accumulation algorithms and cell size. Our cell size-flow accumulation analysis (Figure 7) suggested that the performance of multiple flow direction algorithms did in fact decline inversely with cell size. Unlike Park et al. (2009) they did not converge with single flow direction algorithms. This is likely due to the fact that the maximum cell size we investigated was 10m.

6) The comparison of TWI and VWC becomes vague in its shortness. The relation of TWI and deviation from mean VWC is not accounted for. A study or a comprehensive presentation of references should be inserted.

RESPONSE: We normalized all VWC readings by the mean VWC at each sampling date to facilitate comparisons across multiple dates and sites. Our main objective was not to investigate the physical links between TWIs and VWC, but to investigate which of the myriad TWI formulations best approximates observed soil moisture patterns. We agree that direct linkages between TWI and VWC would be a valuable analysis and allow greater consideration of the physical hydrology, but is somewhat outside the scope of this study and its addition would make an already lengthy paper unwieldy. However, the connection is not totally omitted – for example, please see lines 692-712; Figures 11 & 12, lines 717-738, lines 742-753, lines 755-767, and lines 81-110 and the following references: Burt and Butcher, 1985; Moore et al., 1988; Ladson and Moore, 1992; Jordan, 1994; Schmidt and Persson, 2003; Western et al., 2004; Tague et al., 2010; Nyberg, 1996; Hellstrand, 2012; Beven and Kirby, 1979; Western et al., 1999. If the reviewer knows of important studies we have overlooked, by all means provide citations for us.

7) Table numbering must be edited.

RESPONSE: Corrections made.

8) P14057L6-7: “3m TWIs” is a confusing description. Another way of putting it would be “the TWIs generated from the 3m resolution DEM”

RESPONSE: Correction made. Please refer to lines 497-498.