Hydrol. Earth Syst. Sci. Discuss., 10, C361–C371, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C361/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Regionalization of patterns of flow intermittence from gauging station records" *by* T. H. Snelder et al.

Anonymous Referee #1

Received and published: 15 March 2013

General Comments: The manuscript describes two statistical models that use environmental variables (catchment characteristics, climate, and geology) for predicting the membership of stream segments to flow classes. The flow classes were defined based on the duration and frequency of zero-flow or drying according to hydrological data from gauging stations throughout France. The intent of the first model was to predict whether or not segments are perennial or intermittent developed hydrological data (23-35 y) from 628 gauging stations. The intent of the second model was to discriminate among three intermittent classes that were based roughly on the even distribution of the mean annual frequency of zero-flow periods and the mean duration of zero-flow periods across 123 intermittent gauging stations. The authors reported that the statistical models using the environmental variables did a fair to poor job of classifying segments correctly (i.e., match classifications based on gauging data assignments). The authors

then use the models to map the probability of intermittence for ca. 115,000 river segments across France.

Overall, I found the paper to address an important need and the authors used an approach (Random Forests) that is fairly novel to riverine science especially for characterizing temporary rivers. Although the finding that streams in drier and warmer regions are more likely to have zero-flow periods than those in wetter and cooler regions is not surprising, the authors do take some steps to apply their findings in such a way as to better inform policy and management of rivers in France. The approach and methods used here should work elsewhere (if gauging stations and data are sufficiently available across the perennial-intermittent gradient). Here I outline some of my larger concerns and suggestions and then go on to provide my specific comments.

1. Did runoff or flow characteristics (intermittence, degree of intermittence) have any control over the distribution of gauging stations in France? If so, how might that affect the predictions for the ungauged portions of the network? Were locations of gauges in the network variable (upstream, middle or downstream end of segments) or did they tend to occur near the mouth of catchments (downstream end of segments, near confluences)? In other words, could drying frequency and duration be related to proximity to confluences and therefore influence how representative their data are for segments?

2. Regional weather data used in the model were from 1961-1990. Why not for a longer period of record (through to 2009)? Is it possible that drought conditions (unusually dry and warm) were associated with zero flows recorded at some gauges (Subclass 1, perhaps some of Subclass 2) and these reaches might really be perennial reaches that were prone to drying under drought conditions? The definitions of perennial and intermittent reaches or segments used by the authors should be stated at the start of the introduction. Most definitions of perennial streams characterize them as having year-round flow except in periods of drought. The authors need to indicate to the readers why they did not consider drought in classifying streams (or climatic measures that might be useful in associating zero-flows to drought over longer time scales, rather

than average annual rain and temperature for climatic variables).

3. Related to #2, the division of the intermittent segments into subclasses appears to be based on evenly distributing the stations by DUR and FREQ data as opposed to possibly more meaningful divisions (functional or policy or management related). The distribution of stations appears to be skewed toward more perennial waters (most to left of 0.5 mFreq in Fig.4 and median DUR only 7.3 days) and away from what authors later describe (P1530 L16-18) as channels permanently above the water table (i.e., ephemeral). The authors characterize two different types of classes in the discussion (P1530 L12-18) based on the channel and groundwater table elevations (and therefore predominant sources of flow) that would probably be a more functional way to classify temporary streams. Are ephemeral or episodic streams (sensu Williams 2006 Biology of Temporary Waters) lacking or absent in most regions of France or are they common but lack stream gauges? How do the authors foresee their 3 subclasses and their associated boundaries being useful or meaningful to those who might use these maps? (e.g., distinction between 4 days versus 6 days of zero-flow per year on average seems rather arbitrary). My point related to #2 above is if gauges were more likely positioned in segments with perennial flow (or nearly perennial flow) than in segments with infrequent and short durations of flow then the authors should reconsider how to classify/interpret their dataset. It might be more meaningful especially from a policy or management perspective, if in fact some of the stations the authors identified as intermittent Subclass 1 and/or 2 are really perennial sites that under drought conditions (or perhaps abstraction) are particularly susceptible to having zero-flows.

If the authors are attaching significance to the subclasses it would seem useful for readers to know where across the two gradients or on the biplot (Fig.4) the misclassifications were most prevalent. I suspect that they were near the boarders, but may only be restricted to the intersection of the 3 subclasses or predominant along 1 boundary.

4. The focus on of the paper is on the discrimination of perennial from intermittent and among subclasses of intermittent segments and their relationship to environmen-

tal variables. Which is fine, but the authors do not consider in their analysis or interpretation of results that stations may also vary independently by other hydrological aspects or indices (perhaps more so than zero-flow duration and frequency) and those indices could have strong(er) associations with environmental variables. This is somewhat surprising because some of the authors have used largely the same dataset to develop models to more broadly classify river segments based on largely the same environmental variables.

5. Another consideration for the authors is to apply their perennial-intermittent model for station catchments that were discarded because of various modifications due to reservoirs, diversions, abstractions, etc. Does the predicted classification by the model agree that the assumption that the modifications were severe enough to switch the flow regimes of any of these stations? If so, any patterns regarding the type or magnitude of modification?

Specific comments:

P1512 L10-11: Annually? Over the entire period of record (35 yrs)? Seems like one (or a few) dry period over 35 years might be so infrequent that the stream could be considered perennial and zero flow may be caused by severe drought conditions?

P1514 L6: “consumptive use” implies non-sustainable use. Is the intent by the authors that specific? If not, suggest a different description (e.g., human use).

P1514 L15: This statement really applies only to US federal regulations (i.e., Clean Water Act). State regulations can be more protective of their intermittent streams. Change “few” to “fewer.”

P1515 L26-27: One study that I am aware of that the authors overlooked that uses models (though not using RF or related techniques) with environmental variables and that are developed with gauge data to extrapolate intermittent-perennial membership to ungauged locations is Wood et al. 2009 Estimating locations of perennial streams in

Idaho using a generalized least-squares regression model of 7-day, 2-year low flows. SIR2009-5015 for various regions in the state of Idaho in the US.

P1516 L20-25: Briefly describe the distribution of the gauges and the reason for their placement across the network. Did intermittence have any bearing on the distribution of the gauges in the network? How were the gauges that the authors excluded because of modifications (reservoirs, diversions, abstractions) distributed in the network? (3800 total gauges and only 628 used here and 763 used by Snelder et al. 2009) Were areas related to environmental variables used in the models? Would the distribution of modified flow regimes potentially affect the uncertainty of model predictions for certain HER regions compared to others? If so, on any these, fronts then the authors should account for these in their analyses and/or interpretations/explanations.

P1516 L22-24: Were recording intervals similar across stations that were used to produce daily mean discharge? May not apply to gauging stations in this study, but flow events <1 day on ephemeral drainages occur and if intervals are infrequent these might be missed and affect FREQ and DUR.

P1517 L1: Provide an explanation of why the authors chose 35 years (but as low as 23 years) for this study but a minimum of 20 years (over 30 year period) for Snelder et al 2009. (628 stations here vs 763 in Snelder et al. 2009). Does having 21% fewer stations with longer periods of records improve differentiating intermittent versus perennial segments? Seems like 20 years should be a sufficient timeframe for characterizing segments as being either intermittent or perennial, but longer periods might improve documenting perennial sites that may be prone to drying under drought conditions (subclass 1 and maybe subclass 2).

P1517 L2-3: Were there stations with years having multiple gaps <20 days? Provide the range for cumulative number of missing days for a year. Please indicate whether or not there were there any gaps <20 day long that occurred immediately prior to or after a zero-flow period? If there were, how were these handled in terms of determin-

ing DUR? How might this influence certainty in classifying (and therefore predicting) among the intermittent subclasses? Indicate whether the authors used the calendar year (Jan-Dec) or hydrologic year (Oct-Sept) or some other separation between years in organizing the data. Please indicate whether there were any drying events that extended from one year to the next and explain how these were handled in calculating FREQ and DUR for those consecutive years sharing the same zero-flow period (and thereby affecting mFREQ and mDUR).

P1517 L19-21: Was 2.5 km2 the drainage area of the smallest gauged catchment? If so, please indicate that here or explain why the authors chose this as the minimum catchment size.

P1518 L14-21: Was the time frame for nDryDays and dDry based over the same timeframe (1961-1990) as the other climatic variables?

P1518 L23-25: Was the 1:250,000 scale maps have sufficient resolution to include channels at all of the gauge locations? Based on the authors' statements on P1514 L24-27, maps of this scale exclude channels. In the US, the 1:250,000 scale maps are consider coarse and exclude many tributary streams (therefore much of the channel length), especially those with smaller catchment areas. Using maps of this scale should underestimate the actual drainage density, but the degree might vary among regions or networks (underestimate drainage density for "round" networks more so than for "long" networks).

P1519 L5-12: Please clarify/specify whether or not Hard and Perm values were weighted based on catchment surface area or some other means.

P1521 L14-15: The DEM-based or the river channel map network?

P1526 L16-19: While I think these partial dependence plots are very useful for readers to interpret the variables' relationships to the classifications, how does one objectively separate these types of responses into these 3 types (increasing, decreasing, and U)?

To me, Perm response (identified as U) looks a lot like Rain response (identified as decreasing) and the WinSumRain response (identified as U) looks like the Tmax response (increasing) or the Rain response (decreasing). Are these supposed to describe only the pattern between the rug marks along the x-axes? If so, then the patterns might be more clearly differentiated if the authors limit the response line between the deciles.

P1526 L19-22: Would the low and evenly weighted importance measures suggest that there was also little difference between the reduced model retained and other models?

P1526 L22-24: Drainage density also insignificant?

P1526 L26: DEM-based?

P1526 L25-27: So does the legend in Fig. 8 reflect the probability thresholds in the along the x-axis of the right panel of Fig. 6 and the 39% estimate is based on concluding that all of the segments coded in Fig. 8 with probabilities greater than or equal to 0.35 are intermittent and those less than 0.35 are perennial? Also does this 39% estimate exclude portions of river network with modifications that the authors chose to exclude from the station dataset? If not, then do the authors think a more realistic value would be higher, lower, or about the same (modifications are rare)? Some thing that could be briefly addressed in the discussion.

P1526 L28: for consistency suggest capitalizing Hydro-Ecoregions (as on P1516 L9).

P1527 L13: Should be Fig. 9b (no Fig. 8b)?

P1527 L20: Maybe cite Fig. 7 here as well (to show relationships of these variables with intermittence)?

P1528 L4-7: Drainage density and SumWinRain also insignificant? Interesting that these were among the best 3 predictors for Snelder et al.'s (2009) flow class 6 (representing intermittent stations), but are not useful predictors here (and drainage density apparently wasn't useful to separate perennial and intermittent). This seems like something that would be relevant for discussion but was overlooked by the authors.

P1529: These discussion paragraphs largely restate results and provide little additional insight than what was already stated earlier.

P1529 L3-5: Suggest inserting "some" between "were" and "significant" because not all geology and climatic variables were useful in predicting intermittence.

P1529 L25-29: As opposed to what for Subclass 3? This information is again restated on P1530 L12-15.

P1530 L1-7: Any snowpack in these regions with steep slopes (eastern France) to supply flows?

P1530 L20-27: This points to the possibility of drought (climatic variation at varying temporal scales) influencing the weak spatial synchronization. If there are droughts that do not occur simultaneously over the entire country, wouldn't one expect that all stations would exhibit the same temporal pattern? Frequency is not always a clear measure of intermittence. An extremely dry year or ephemeral stream may result in just couple dry event that may last very long durations and interupted by a single short flow period, whereas a less extreme dry year or stream with periodic connections to groundwater may have several drying events of short durations.

P1531 L5: How important were the hydrologic indices for zero-flow in discriminating flow classes in Snelder et al. (using the same gauging stations) relative to the other kinds of hydrologic indices? How well does Snelder et al (2009) flow regime class 6 ("intermittent-flashy regime") align the authors' assignments for stations in present study and with intermittence probabilities on Fig. 8 in this manuscript?

P1531 L9-11: This was not already reported/interpreted in the Snelder et al. (2009) paper?

P1531 L29: Perhaps the ordinal assignment of geology make this a coarser predictor variable also? More detail on thickness and associated soil information might improve the predictive power?

P1532 L1-5: The authors' language suggests a certainty that smaller-scale factors are needed to improve the predictions. While I do not wholly disagree with this statement, how it is stated seems a bit strong considering no data is presented here to support it. Another consideration for the fair to poor performance of the models developed in this study that the authors do not address is that stations could vary more in other aspects of their hydrology (many detailed in Snelder et al 2009) than just DUR and FREQ and these other aspects or indices have associations with the environmental variables used in this study. This seems to be supported by what is presented in Table 3 of Snelder et al. 2009. Compare, for instance, flow classes 6 and 3, both of which based on Figure 3 have stations with periods when mean daily flows are zero. The PCA axes centroids for these two classes varied across other PCA axes than just PCA axis 9 (only accounted for 3% of the explained variation across all stations) which had the strongest correlation to DUR and/or PCA axis 6 (also only accounted for 3% of explained variation) which had the strongest correlation to frequency of low flows.

P1532 L18: Clarify please what is meant by "reconfiguring the gauging network." Do the authors here saying to move gauges from perennial to intermittent segments? Locating gauges randomly or probabilistically throughout the network?

P1532 L26-28: Particularly subclass 3 intermittent segments.

P1533 L3-5: Depends on the HER, correct? The drier and warmer HERs have higher error and a tendency to over predict gauges are intermittent according to Figure 9. Why didn't the authors include HER class as a predictor variable in the RF? Seems like that might be important predictor and support this statement.

P1533 L7-8: see also Ademollo et al. 2011 Trends in Analytical Chemistry 30:1222-1232

(did not carefully edit the references, these are just a couple typos that I happened to notice) P1534 L7: Author name spelled differently than on P1533 L12.

P1537 L18: Should be "in the arid Negev"

Table 1: Are these values for drainage density and Shape correct? How does one have a drainage density or Shape of zero? Is this because of the map scale used? Or are the values very small and the authors chose to round down to zero? If the later, then maybe show as e.g., >0.0001. What about zero values for Hard and Perm? Do catchments with zero for these variables not have any of the geological categories listed in Table 2?

Table 1: Check descriptions for Chalk and Lime.

Figure 1: spell out HER

Figure 2 (legend and heading): throughout main text these are called intermittence subclasses.

Figure 3: y-axis labels move to left side for FREQ.

Figure 4: Would be useful to identify the stations that were misclassified by the flow-intermittence model (circle symbols or use inset). Could also identify those intermittent stations that the flow-regime classification model misclassified or at least indicate in the text how misclassifications were distributed among the 3 subclasses.

Figure 6: Spellout/define ROC and PCC in heading. Describe what the black circles represents in the right panel.

Figure 8: Legend includes one bin for "0.3 – 0.3" Is this supposed to represent the 0.35 probability threshold from which the 39% was derived? If so, please add explanation to the figure heading (maybe label as "0.35 threshold" in legend) ; otherwise delete this bin from the map and legend.

Figure 9: Capitalize Hydro-Ecoregions in figure heading for consistency and follow with abbreviation in parentheses (later in heading referred to by HER). Maybe consider referencing Figure 1 in this heading. How balanced (relative to their area) are the

stations across the various HERs? It requires the readers to look between Fig 1 and Fig 2 to give some sense of the distribution. If one considers HER 13 (Landes) it doesn't appear that there are as many stations in this HER as some of the others. So only 2 intermittent stations across only 8 or 9 total stations could make the proportion of intermittent gauges appear higher than it might actually be had more stations been located in that HER. Showing the total number of stations for each HER will help the readers more fully interpret Fig 9. Perhaps an efficient way to do this would be putting the number of stations within each HER in parentheses in the legend of Figure 1.

Figure 10: In the main text these are referred as 3 "subclasses."

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 1511, 2013.