# Local Quasilikelihood in Catchment Modification Detection

G. Capkun[a,b], A.C. Davison[a] and A. Musy[b]

[a] *Department of Mathematics, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland (gorana.capkun@epfl.ch, anthony.davison@epfl.ch)*

[b] *Department of Rural Engineering, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland (andre.musy@epfl.ch)*

**Abstract:** The hydrological literature contains three basic approaches dealing with impact of catchment modifications on runoff: the experimental catchment approach, the modification modelling approach, and studies involving the use of hydrological models. In this paper we propose an extension of the model estimate and dynamic response variable comparison methods (before and after changes). A robust quasilikelihood for joint modelling of mean and variance is fitted locally in order to give smooth estimates of the time-varying parameters. We give an example using the Viège catchment in Switzerland where the Mattmark dam was constructed in 1960's. The correlated series of the local parameter estimates detected this change in catchment response.

***Keywords:*** Catchment modifications; Local modelling; Quasilikelihood; Time-varying parameters

## 1. INTRODUCTION

Rapidly increasing population pressure in many rural areas has led to changes in land use owing to deforestation, reclamation of wetlands, urbanization, etc., or due to other catchment modifications like dams and roads. Such changes are intended to increase agricultural production, the use of water power, improve the quality of life and so forth. However, land mismanagement may have inadvertent negative effects on a hydrological regime, such as increasing the occurrence of floods and decreasing dry season flows. Thus there is a need for improved knowledge and quantitative analysis of the impact of these changes and management practice on land and water resources.

The hydrological literature contains three basic approaches dealing with impact of catchment modifications on runoff: the experimental catchment approach [Cerdà, 1998], the modification modelling approach [Iritz, et al., 1994], and studies involving the use of hydrological models (sometimes combined with basic statistical methods such as linear regression and simple parametric and nonparametric tests) [Lørup, et al., 1988].

This paper describes a simple method of detecting changes in catchment response that uses a hydrological rainfall-runoff transfer model as its basis. A simple technique proposed is applied to a quasilikelihood-based model [Capkun, et al., 2001], but can also be used in combination with other hydrological rainfall-runoff models.

## 2. TIME-VARYING PARAMETERS

### 2.1 Introduction

Studies involving hydrological models in catchment modification detection are based on sensitivity analysis to illustrate the hydrological response to various conservation practices. Lumped catchment models must be used carefully as they may fail to predict the impact of the change on catchment runoff due to limitations in the model conceptualization of the hydrological processes involved [Kuczera et al., 1993]. Therefore, rigorous model validation procedures are required before the model capabilities can be assessed [Ewen and Parkin, 1996].

In our previous work [Capkun, et al., 2001], we have developed a simple Markov generalized linear model of the mean and variance structures of runoff at time $t$, given previous rainfall and runoff. Its mean is taken to be a linear autoregressive combination of present and previous rainfall and previous runoff, while its variance also depends on rainfall history. A more formal description of it is given in Section 2.2. Inference for its parameters may be performed using classical likelihood methods, and also using the more robust technique of quasilikelihood, presupposing no particular distribution for runoff. Robust "sandwich" confidence intervals for the model parameters are constructed using both likelihood and quasilikelihood approaches in order to account better for model uncertainty. The model was fitted and validated on 13 Swiss Plateau catchments, where it showed good fitting and short-term predicting characteristics. We therefore judge it to be suitable for detecting hydrologic-changes in a catchment.

In order to compare the periods before and after catchment modifications, it is necessary to know when such a change happened, and which period can be considered as transition between the "before" and "after" change. In order to detect the existence of the land use changes in the Balquhidder catchments, Jakeman, et al. [1993] fit a simple unit hydrograph-based model for each of 5 years of measurements separately. They conclude that no land use changes can be detected because of the high interannual variability.

Our proposition below can be viewed as the extension of Jakeman's model estimate and dynamic response variable comparison methods (before and after changes) and is described in Section 2.3.

## 2.2 Quasilikelihood Model

Here we first give a brief description of the rainfall-runoff transfer model that is used in this paper. We suppose that the runoff at time $t$, $Y_t$ given the present and past rainfall $X$ and past runoff

$$H_t = (X_t, X_{t-1}, \ldots, X_{t-k}, Y_{t-1}, \ldots, Y_{t-l}),$$

comes from some unknown density with mean $\mu_t$ and variance $\sigma_t^2$ . The rainfall is taken to be known and the model can now be written in the following form

$$\mu_t = E(Y_t | H_t) = \sum_{i=0}^{k} \beta_i x_{t-i} + \sum_{j=1}^{l} \gamma_j y_{t-j},$$

$$\sigma_t^2 = \text{var}(Y_t | H_t) = \frac{1}{\exp(\lambda_1 + \lambda_2 I_{t,k})} \mu_t^2,$$

$$I_{t,k} = \begin{cases} 0, & \text{if } x_{t-1} = \cdots = x_{t-k} = 0, \\ 1, & \text{otherwise.} \end{cases}$$

$$t = r+1, \ldots, T; \quad r = \max(k, l),$$

where $k$ and $l$ need to be found in the model choice procedure. The model parameters

$$\theta = (\beta_0, \ldots, \beta_k, \gamma_1, \ldots, \gamma_l, \lambda_1, \lambda_2),$$

are taken to be time-invariant and are estimated using one version of iteratively reweighted least squares algorithm [Green, 1984], GEN-IRLS.

## 2.3 Local Modelling

If the calibration period is too long, say 30 years, we can imagine that due to the changes in catchment response, the parameter estimates that quantify the catchment reaction would also change. The consequence of fitting a "constant parameter" model described above on the whole series would result in poor diagnostics (e.g. large residuals). The reason for that is our will to "impose" the same behaviour to the whole data series when maximizing a single overall objective function.

Instead of having a constant vector parameter $\theta$, we consider instead a vector function of time

$$\theta(t) = (\beta_0(t), \ldots, \beta_k(t), \gamma_1(t), \ldots, \gamma_l(t), \lambda_1(t), \lambda_2(t)).$$

For a given time $t$ the new model with time-varying coefficients reduces to the constant-parameter model given above.

To estimate $\theta(t)$ we suppose that given $H_t$, all the $Y_t$ belonging to a time period of certain width centred at time $t$ share the same constant true value of parameter $\theta$. By sliding a time window of size width over the observations only the newest observations are seen and used for a constant parameter estimation using the GEN-IRLS algorithm mentioned above. At each sliding, a new parameter estimate $\theta_s$, $s=1,\ldots,n$ is obtained. The confidence bounds around the estimated curves for the parameters of the mean $\mu_t$

$$\widehat{\beta}_0(t), \ldots, \widehat{\beta}_k(t), \widehat{\gamma}_1(t), \ldots, \widehat{\gamma}_l(t),$$

can be calculated [Capkun, et al., 2001]. The moving window allows the model to adapt to the variations in system dynamics. Depending on how quick the dynamics change, we slide our data window by a bigger or smaller step number of observations; step $\leq$ width. The final result is a smooth nonparametric estimate of $\theta(t)$;

$$\hat{\theta}_1, \ldots, \hat{\theta}_n,$$

where $n = n$(width, step) is a total number of moving windows. Table 1 illustrates the method described above. It is worth noting that the approach by local modelling is adequate when the model parameters

**Table 1.** Local modelling scheme: `width` = 5, `step` = 2. First row: the whole data series and its parameter estimate. Other rows: data from the window are used to calculate the corresponding estimate $\hat{\theta}_s$.

| Data | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Data | $X_1$<br>$Y_1$ | $X_2$<br>$Y_2$ | $X_3$<br>$Y_3$ | $X_4$<br>$Y_4$ | $X_5$<br>$Y_5$ | $X_6$<br>$Y_6$ | $X_7$<br>$Y_7$ | $X_8$<br>$Y_8$ | $X_9$<br>$Y_9$ | $X_{10}$<br>$Y_{10}$ | $X_{11}$<br>$Y_{11}$ | $X_{12}$<br>$Y_{12}$ | $\hat{\theta}$ |
| Window 1 | $X_1$<br>$Y_1$ | $X_2$<br>$Y_2$ | $X_3$<br>$Y_3$ | $X_4$<br>$Y_4$ | $X_5$<br>$Y_5$ | | | | | | | | $\hat{\theta}_1$ |
| Window 2 | | | $X_3$<br>$Y_3$ | $X_4$<br>$Y_4$ | $X_5$<br>$Y_5$ | $X_6$<br>$Y_6$ | $X_7$<br>$Y_7$ | | | | | | $\hat{\theta}_2$ |
| ⋮ | | | | | | | | | | | | | ⋮ |
| Window n | | | | | | | | | $X_9$<br>$Y_9$ | $X_{10}$<br>$Y_{10}$ | $X_{11}$<br>$Y_{11}$ | $X_{12}$<br>$Y_{12}$ | $\hat{\theta}_n$ |

are functions of the same explanatory variable; time in our case. If we suppose that the parameter values depend on individual explanatory variables, estimation methods for additive models should be used [see for example Hastie and Tibshirani, 1993].

If we expect the time variation to be fast, the `width` should be small, otherwise the estimates could be seriously biased. However, fast adaptation means that a smaller portion of the data are used for the estimation, which results in noisy estimate. Therefore, the choice of `width` leads to a bias/variance trade off. In practice, the value of `width` would depend on the information we would like to extract from the data series; if we take `width`=3 months the seasonality of the catchment reaction is observed, while for `width`=1 year, our goal is to detect changes that are slow in time. Taking `width` to be the length of the whole series reduces the local approach to a constant parameter model. Jakeman's approach is the special case when setting `width`=`step`.

The correlation in the series of estimated parameters depends on the ratio between `step` and `width`; the degree of dependence is higher for values of the ratio close to zero.

## 3. APPLICATION

### 3.1 Viège Data

Viège is an alpine catchment situated in Switzerland with a surface of 780 km$^2$ and its outlet in Viège. Its mean altitude is 2800 meters with a glacier proportion of 30% [AHS, 2000]. It is classified as "b-glaciaire" type of catchment by its hydrological regime [Weingartner and Aschwanden, 1992], with a maximal monthly mean runoff in the summer season (June - August). The input of the rainfall-runoff system in summer is mainly composed of precipitation, snowmelt and glacier melt. During the winter period the precipitation is stored as snow which explains very low winter runoff measurements.

In 1960's, the Mattmark dam and a deviation of one part of the streamflow towards the Grande-Dixence dam were constructed. These modifications allowed for the exploitation of the hydroelectric potential of the catchment and considerably modified its hydrological regime.

Daily rainfall and runoff data for the period of 1922 -1998 are shown in Figure 1. The rainfall data are measured at the Zermatt rainfall station [identification number 7190 in Weingartner and Auschwanden, 1992]. The mean behaviour of the rainfall series seems stationary over time, but there are more extreme rainfall events in recent years. A considerable reduction in the mean runoff regime is observed after the dam construction.

Viège catchment is a "textbook example" of a changed regime and we hope that the local modelling technique proposed in Section 2.3 will detect this change in catchment response.

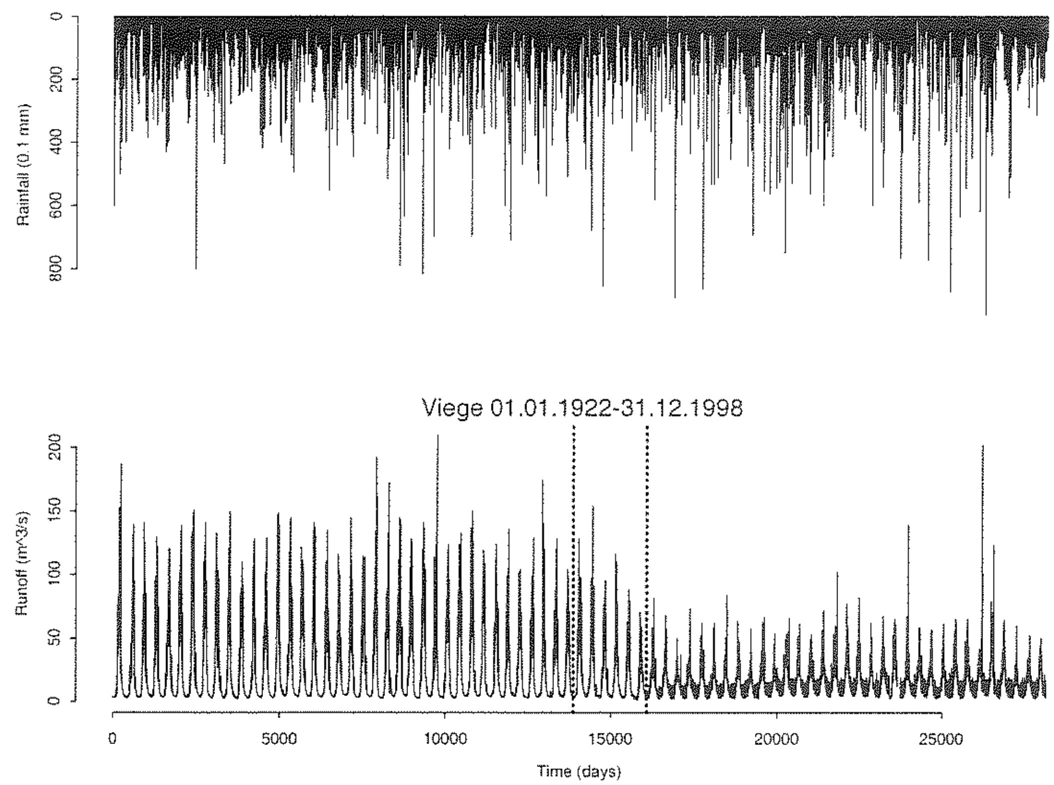


**Figure 1.** Daily measurements of Viège data (01.01.1922 – 31.12.1998). The top panel shows the rainfall data in 0.1 mm measured at the Zermatt rainfall station. The bottom panel shows the runoff data in m$^3$/s at the Viège outlet. Vertical dotted lines show the approximate beginning and end of the dam.

## 3.2 Results

In order to choose the appropriate model for the local fitting, the first four years of Viège data were selected. The model choice procedure proposed in Capkun, et al. [2001] suggests a model of form

$$\mu_t = \beta_1(t)x_{t-1} + \gamma_1(t)y_{t-1} + \gamma_2(t)y_{t-2},$$

$$\sigma_t^2 = \frac{1}{\exp(\lambda_1(t) + \lambda_2(t)I_{t,k})}\mu_t^2.$$

The width is taken to be 4 years (1460 daily observations) with a step of 1 month (30 daily observations). Local estimates of 5 model parameters are plotted over time in Figure 2. Two solid vertical lines stand for the approximate beginning (1960) and the end (1965) of dam construction. The same pattern is observed in all estimated parameter series: rather stationary behaviour before the beginning of construction, and "transfer" nonstationarity during the construction. The estimated series for the parameter $\beta_1$ (the top left panel of Figure 2) has a larger mean and variance after the dam construction; this parameter explains the direct impact of the rainfall at time *t-1* on the runoff at time *t*. The autoregressive parameters $\gamma_1$ and $\gamma_2$ (middle panel in Figure 2) are equally well estimated for the whole series of data (the variance remains stable). A nonstationary trend observed after the dam construction may be related to the working regime of the dam; the data needed to confirm this are unfortunately not available. The estimates of the variance parameters $\lambda_1$ and $\lambda_1+\lambda_2$ diminish after the construction, making the dispersion parameter

$$\frac{1}{\exp(\lambda_1(t) + \lambda_2(t)I_{t,k})},$$

and of the runoff larger than before the construction. This may be explained by the noisy runoff measurements which are due to the working regime.

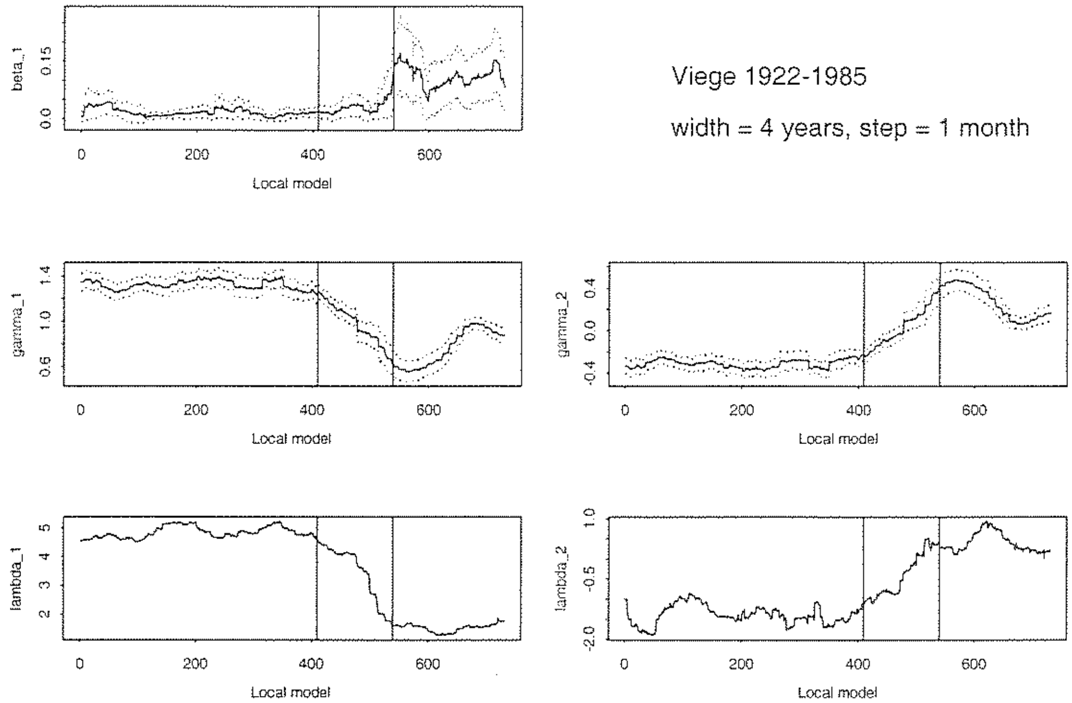


**Figure 2.** Local parameter estimates (solid line) for Viège daily data; `width` = 4 years (1460 days), `step` = 1 month (30 days). Vertical lines show the approximate beginning and end of the dam; dashed lines are "sandwich" confidence intervals [Capkun, et al., 2001].

## 4. CONCLUSIONS

In this paper we propose a simple methodology for local estimation of time-varying parameters, which is simple and easy to program and does not depend on the rainfall-runoff model. The appropriate choice of the time window, `width`, can help us to extract different information from the data series, such as the nonstationarity due to climatic, catchment or land use changes, unusual behaviour, etc.

This approach seems not to have been applied *per se* in the hydrological literature, which apart from applications of smoothing methods to financial times series has usually concentrated on the situation where the data are independent though not identically distributed [Fan and Gijbels, 1996]. Not many smoothing methods have been used in hydrology, and if so, almost entirely in the density estimation context. However they seem potentially very useful [Ramesh and Davison, 2001].

## 5. ACKNOWLEDGEMENTS

We acknowledge the support of the Swiss National Science Foundation, project FN 2100-049614.

## 6. REFERENCES

AHS, Annuaire hydrologique de la Suisse, Service Hydrologique et Géologique National, Office Fédéral de l'Environnement, des Forets et du Paysage, 2000.

Capkun, G., A.C. Davison, and A. Musy, A robust rainfall-runoff transfer model, Accepted for publication, *Water Resources Research*, 2001.

Cerdà A., Changes in overland flow and infiltration after a rangeland fire in a Mediterranean scrubland, *Hydrological Processes*, 12, 1031-1042, 1998.

Ewen, J., and G. Parkin, Validation of catchment models for predicting land-use and climate change impacts, 1. Method, *Journal of Hydrology*, 175, 583-594, 1996.

Fan, J., and I. Gijbels, *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London, 1996.

Green, P.J., Iteratively reweighted least squares for maximum likelihood estimation, and some robust resistant alternatives (with Discussion), *Journal of the Royal Statistical Society, Series B*, 46, 149-192, 1984.

Hastie, T., and R. Tibshirani, Varying –coefficient Models. *Journal of the Royal Statistical Society, Series B*, 55, 757-796, 1993.

Iritz, L., B. Johansson, and L. Lundin, Impacts of forest drainage on floods, *Hydrological Sciences*, 39, 637-661, 1994.

Jakeman, A.J., I.G. Littlewood, and P.G. Whitehead, An assessment of the dynamic response characteristics of streamflow in the Balquhidder catchments, *Journal of Hydrology*, 145, 337-355,1993.

Kuczera, G., G.P. Raper, N.S. Brah, and M.D. Jayasuriya, Modelling yield changes after strip thinning in a mountain ash catchment: an exercise in catchment model validation, *Journal of Hydrology*, 150, 433-457,1993.

Lørup, J.K., J.C. Refsgaard, and D. Mazvimavi, Assessing the effect of land use change on catchment runoff by combined use of statistical tests and hydrological modelling: Case studies from Zimbabwe, *Journal of Hydrology*, 201, 147-163, 1998.

Ramesh, N.I., and A.C. Davison, Local models for exploratory analysis of hydrological extremes, Under review, *Journal of Hydrology*, 2001.

Weingartner, R., and H. Auschwanden, Hydrological Atlas of Switzerland, Swiss National Hydrological and Geological Survey, Berne, 1992.